

Characterising the binding interactions and thermodynamics of Odour Binding Protein 3

Katherine Louise Portman BSc. (Hons)

Thesis submitted to the University of Nottingham for the
degree of Doctor of Philosophy

July 2012

ABSTRACT

Odour Binding Proteins (OBPs) are found in the olfactory system of a range of species. Whilst invertebrate OBP function is well understood, the exact function of these proteins in the vertebrate nasal mucus is not fully understood. Multiple subtypes of rat OBPs have been identified and found to share less than 30% sequence identity. Studies have suggested each rat OBP binds to particular sets of odours, which may afford them a particularly important role within the olfactory system, pre-sorting odours.

This study focuses on OBP3, closely examining the binding interaction of this protein with a range of odours. This has been done using Isothermal Titration Calorimetry which revealed that the binding of the highest affinity ligands, the heterocyclic compounds, is enthalpically driven. A defined odour series, the γ -lactones showed that despite increasing ligand size and hydrophobicity, the free energy of binding of these ligands is maintained.

Interactions with both 2-isobutylthiazole and the γ -lactones were examined using NMR spectroscopy, which required the NMR assignment of OBP3 to be determined. In addition a homology model of OBP3 was created in order to structurally map the per-residue changes of OBP3 upon binding. It has been found that OBP3 is able to subtly adjust in order to accommodate each of these ligands.

Protein engineering of the OBP3 binding pocket has been used to highlight the importance of its size and hydrophobicity. The importance of a tyrosine residue that appears to cover the opening to the binding pocket and is conserved across both the OBPs and the lipocalins family they are part of, has been demonstrated. Mutagenesis has also revealed the importance of a number of key residues for the binding of 2-isobutylthiazole. The ability to rationally improve the affinity of OBP3 for a particular odour has also been demonstrated.

ACKNOWLEDGEMENTS

The first acknowledgement must go to Dr. David Scott, for introducing me to the subject of Odour Binding Proteins and for his supportive nature and unending enthusiasm, which has continually spurred me on. I also thank Prof. Mark Searle for his encouragement and for many useful discussions throughout my PhD. I am extremely grateful to Dr. Jed Long who has spent a large amount of time teaching and helping me with a number of techniques, particularly NMR. I also thank him for numerous useful ideas and discussions. In addition I would like to thank Dr. Huw Williams for taking care of the NMR and computer facilities and for always being willing to offer support and assistance.

I acknowledge Masayuki Yabuki and Khairi Zainol, thank you both for not only the academic discussions but also for the numerous interesting chats about so many different topics. Also thanks to Ina Tögel and Jeanette Farshadi for their technical assistance with this project. I am thankful to all who made Sutton Bonington a friendly place to work, in particular Dr. Karin Heurlier, Dr. Neil Doherty, Selina Clayton and Emily Boothroyd. Also to members of the Searle group past and present for their support and friendship both in and out of the lab.

I thank my family, particularly my parents, for their absolute belief in me and for their unwavering support which has always allowed me to achieve my ambitions. Finally thank you to Ian, for being my rock, for constantly encouraging me and without whom I am sure this PhD would not have been completed.

CONTENTS

ABSTRACT	1
ACKNOWLEDGEMENTS	3
ABBREVIATIONS.....	8
1. INTRODUCTION	10
1.1 Olfaction	10
1.1.1 Odours	12
1.1.2 The invertebrate olfactory system	13
1.1.3 The vertebrate olfactory system	20
1.2 Possible roles of vertebrate Odour Binding Proteins	26
1.3 The structure of vertebrate Odour Binding Proteins.....	29
1.3.1 Lipocalins	29
1.3.2 Bovine and Porcine OBP.....	34
1.3.3 Rat OBP 1.....	38
1.3 Previous studies of Odour Binding Proteins	40
1.4 Major urinary proteins	44
1.5 Aims and objectives of this study	46
2. MATERIALS AND METHODS.....	47
2.1 Molecular Biology	47
2.1.1 Strains.....	47
2.1.2 Standard growth conditions.....	48
2.1.3 Strain storage.....	48
2.1.4 Production of electrocompetent XL1-Blue <i>E.coli</i>	48
2.1.5 Transformation of electrocompetent XL1-Blue <i>E.coli</i> cells.	49
2.1.6 Transformation of chemically competent M15 <i>E. coli</i> by heat shock	50
2.1.7 Preparation of purified pQE31 plasmid.....	50
2.1.8 Site directed mutagenesis of OBP3	51
2.1.9 Sequencing	54
2.2 Production of Rat OBP3	55
2.2.1 <i>E. coli</i> expression systems used	55
2.2.2 Plasmids.....	55
2.2.3 OBP3 Sequence.....	57

2.2.4 Overexpression of non-isotopically labelled OBP3 in <i>E. coli</i>	58
2.2.5 Overexpression of isotopically labelled OBP3 in <i>E. coli</i>	59
2.2.6 Selective non-isotopic labelling of OBP3	60
2.2.7 Extraction of OBP3	60
2.3 Purification of OBP3.....	61
2.3.1 Metal Affinity Chromatography.....	61
2.3.2 Gel Filtration	61
2.3.3 Endogenous Ligand Removal	62
2.3.4 Desalting.....	62
2.4 Assessment of protein and purity and quantity	63
2.4.1 Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis	63
2.4.2 Mass spectrometry.....	64
2.4.3 Determination of protein extinction coefficients.....	65
2.5 Biophysical and structural techniques	66
2.5.1 Isothermal titration calorimetry experiments	66
2.5.2 Preparation of odorants for ITC experiments.....	68
2.5.2.1 Fitting of Isothermal titration calorimetry data.	69
2.5.3 Nuclear Magnetic Resonance Spectroscopy	70
2.5.4 Crystallisation trials of OBP3.....	73
2.5.5 Circular Dichroism	74
2.5.6 Differential scanning calorimetry.....	76
2.6 Homology Modelling	77
3. OVEREXPRESSION, PURIFICATION AND CHARACTERISATION OF OBP3.....	78
3.1. Introduction	78
3.2 Results and Discussion	78
3.2.1 Overexpression of OBP3	78
3.2.2 Protein purification.....	79
3.2.3 Determination of the molecular weight of OBP3 by mass spectrometry	87
3.2.4 Assessment of the structure of OBP3	88
3.2.5 Assessment of the thermal stability of OBP3	93
3.3 Conclusions.....	97
4. NMR ASSIGNMENT OF OBP3	99

4.1 Introduction.....	99
4.2 Results and Discussion	106
4.2.1 Homonuclear experiments.....	106
4.2.2 Isotopic labelling	108
4.2.3 Heteronuclear Experiments	110
4.2.4 Three Dimensional ^1H / ^{15}N Heteronuclear NMR experiments	113
4.2.5 Triple Resonance NMR experiments	115
4.2.6 NMR experiments to identify amino acid type	118
4.2.7 Assignment of OBP3 bound to 2-isobutylthiazole.....	123
4.3 Conclusions.....	131
5. INITIAL BINDING CHARACTERISATION OF OBP3	133
5.1 Introduction.....	133
5.1.1 Isothermal titration calorimetry	134
5.1.2 Selection of odorants	138
5.2 Results and Discussion	146
5.2.1 Thiazole and pyrazine ITC binding experiments	146
5.2.2 ITC binding of ligands containing a hydroxyl group	153
5.2.3 ITC binding of molecules containing a double bonded oxygen atom	155
5.2.4 Ligand binding by ITC	158
5.2.5 Constructing a homology model of OBP3	159
5.2.6 Binding of 2-isobutylthiazole to OBP3 by heteronuclear NMR spectroscopy	172
5.3 Conclusions.....	184
6. USING DEFINED LIGAND SERIES TO EXPLORE THE BINDING OF OBP3.....	187
6.1 Introduction.....	187
6.1.1 The γ -lactone series	187
6.2 Results and Discussion	190
6.2.1 D ₂ O exchange.....	190
6.2.2 Binding of OBP3 to the γ -lactones by ITC	196
6.2.3 Binding of OBP3 to the γ -lactones by NMR.....	202
6.2.4 Heteronuclear NOE experiments to probe the mobility of OBP3..	212
6.2.5 Binding of the acetate esters by ITC	225

6.3 Conclusions.....	229
7. PROTEIN ENGINEERING OF THE OBP3 BINDING POCKET	231
7.1 Introduction.....	231
7.1.1 The aromatic nature of the OBP3 binding pocket.....	231
7.2 Results and Discussion	236
7.2.1 Production of the aromatic mutants.....	236
7.2.2 ITC of the aromatic mutants with 2-isobutylthiazole.....	238
7.2.3 ITC of the aromatic mutants with the γ -lactones	249
7.2.4 Design and production of specific binding pocket mutants	259
7.2.5 ITC of the specific mutants with the 2-isobutylthiazole	266
7.2.6 ITC of the specific mutants with the acetate esters	271
7.3 Conclusions.....	274
8. DISCUSSION AND FUTURE WORK.....	276
REFERENCES	282
Appendix I Primers used in this study.....	300
Appendix II Ligands used in this study	301
Appendix III NMR acquisition parameters	302
Appendix IV Crystallisation screens.....	304
Appendix V Chemical shifts of unbound 1 mM OBP3	305
Appendix VI Chemical shifts of 2-isobutylthiazole bound OBP3	309
Appendix VII Chemical shifts perturbations of NH resonances of OBP3 upon ligand binding.....	313
Appendix VIII ^1H / ^{15}N Heteronuclear NOE Values	317

ABBREVIATIONS

1-AMA	1-aminoanthracene
APCI	atmospheric pressure chemical ionisation
CD	circular dichroism
CSP	chemical shift perturbation
DMT	4,5-dimethylthiazole
DNA	deoxyribonucleic acid
ESI	electrospray ionisation
GPCR	G-protein coupled receptor
HSQC	heteronuclear single quantum coherence
IBA	isobutyl acetate
IBMP	2-isobutyl 3-methoxypyrazine
IBT	2-isobutylthiazole
IPTG	isopropyl- β -D-thiogalactopyranoside
ITC	isothermal titration calorimetry
kDa	kilodaltons
LB	Luria-Bertoni
MS	mass spectrometry
MUP	major urinary protein
NMR	nuclear magnetic resonance
NOESY	nuclear Overhauser effect spectroscopy
OBP	odour binding protein
OD	optical density
OR	olfactory receptor
ORN	olfactory receptor neuron
PAGE	polyacrylamide gel electrophoresis
PBP	pheromone binding protein

PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PDB	protein data bank
SBT	2-sec-butyl-4,5-dihydrothiazole
SDS	sodium dodecyl sulphate
TBA	tert butyl acetate
TEMED	N,N,N',N'-tetramethylethylenediamine
TOCSY	total correlation spectroscopy
TOF	time-of-flight
TROSY	Transverse relaxation-optimised spectroscopy
UV	ultraviolet

1. INTRODUCTION

Rat Odour Binding Protein 3 (OBP3) which is the focus of this study is found in the olfactory system of *Rattus norvegicus*. It is one of three subtypes of Odour Binding Protein (OBP) identified in this mammal (Löbel et al. 1998; Löbel et al. 2001). OBPs have been identified in both vertebrate and invertebrate olfactory systems (Pelosi 1994), with the two types of OBPs having distinctly different sequences and 3D structures. The invertebrate OBPs are composed of between six and eight α -helices, connected by flexible linker regions, arranged into a loose bundle (Pelosi et al. 2006). Vertebrate OBPs have anti-parallel β -barrel structures and are members of the larger lipocalin superfamily (Flower et al. 2000). The olfactory systems in which the vertebrate and invertebrate OBPs are found are also very different. The evolution of OBPs as part of both of these divergent systems indicates the likely importance of these proteins.

1.1 Olfaction

Olfaction, or the sense of “smell”, is a chemoreceptive sense, meaning it is used to detect and process chemical signals. The other major type of chemoreception in vertebrates is taste. Olfaction plays a hugely important role in the ability of an organism to process volatile chemicals in its environment. The ability to perceive and discriminate odours is crucial to many species which use these volatile compounds to detect food sources, modulate social

behaviours, identify predators and avoid toxic and potentially infectious materials (Zarzo, 2007)

Interest in human olfaction is driven by the impact odours have on the quality of everyday life. This leads to large commercial interest, for example in fragrance manufacture and also in the food industry. Olfaction is crucial to the sensory perception of food, where, although the sense of taste is important, the sense of smell is vital to the evaluation of flavour (Zarzo 2007; Breer 2008).

There is also commercial interest in invertebrate (particularly insect) olfactory and pheromonal systems. Here understanding can be harnessed in the design of, for example, insect repellents. This would be commercially beneficial from the sale of topological products as well as the protection of crops from parasites (Carey & Carlson, 2011). The distinction between an odour and a pheromone is that an odour (for a terrestrial animal) is a volatile compound that stimulates the sense of smell whilst a pheromone is a chemical substance (often small and volatile) that is used for conspecific (between members of the same species) communication (Kaupp, 2010).

In all organisms which possess chemosensory ability there are specialised cells which are able to process chemicals found in the surrounding environment and relay this information through the nervous system to the brain. In the case of the vertebrate olfactory system these cells are located in the nasal cavity

(Firestein, 2001), whereas invertebrate olfactory cells are found in antennae (Hansson, 2002). Aquatic organisms, such as fish utilise an “accessory olfaction system”, which is able to detect water soluble odorants which diffuse through the skin (Niimura & Nei, 2005; Hamdani & Døving, 2007).

1.1.1 Odours

Although the broad definition of an odour requires that it stimulates the sense of smell, to do this a chemical must possess a number of properties; it must have a degree of water solubility, have high enough vapour pressure (be volatile), low polarity, and be quite lipophilic (able to dissolve in fat). Additionally it must have a relatively low molecular weight. No odour has been found with a molecular weight above 300 Da (Ohloff, 1994). These properties are essential as an odour must be able to travel to the olfactory system through gaseous air, and then, in vertebrates, enter the nasal mucous, which is considerably viscous. This also means that the sense of smell is the most long range of the senses, with odours being able to travel many miles and having the ability to persist for a long time. The sharp cut-off in the molecular size of odours may be due to the fact that vapour pressures decrease rapidly with molecular size, although some chemicals above 300 Da do have appreciable partial pressures but remain odourless, for example the larger musk odours such as tonalide for which many people are anosmic (unable to smell) (Turin, 1996). It has been suggested that they may be too large to fit into the olfactory receptors (Turin & Yoshii, 2003).

1.1.2 The invertebrate olfactory system

As mentioned previously the olfactory cells of insects (the most studied invertebrate olfactory systems) are located in external antennae. The antennae are large in order to accommodate a large number of olfactory sensilla. The sensilla each contain a maximum of four olfactory receptor neurons (ORNs). There are three types of sensilla, each with a different morphology; trichoid (detect pheromones), basiconic (detect food odours and CO₂) and coeloconic (detect food odours, alcohols, amines and water vapour). In addition the maxillary palp in insects such as flies has approximately 60 basiconic sensilla each containing two ORNs. Figure 1.1 shows the basic layout and position of the olfactory system, demonstrating where the antennae (top panel) and sensilla (middle panel) are located. The bottom panel of figure 1.1 shows the layout within a sensillum.

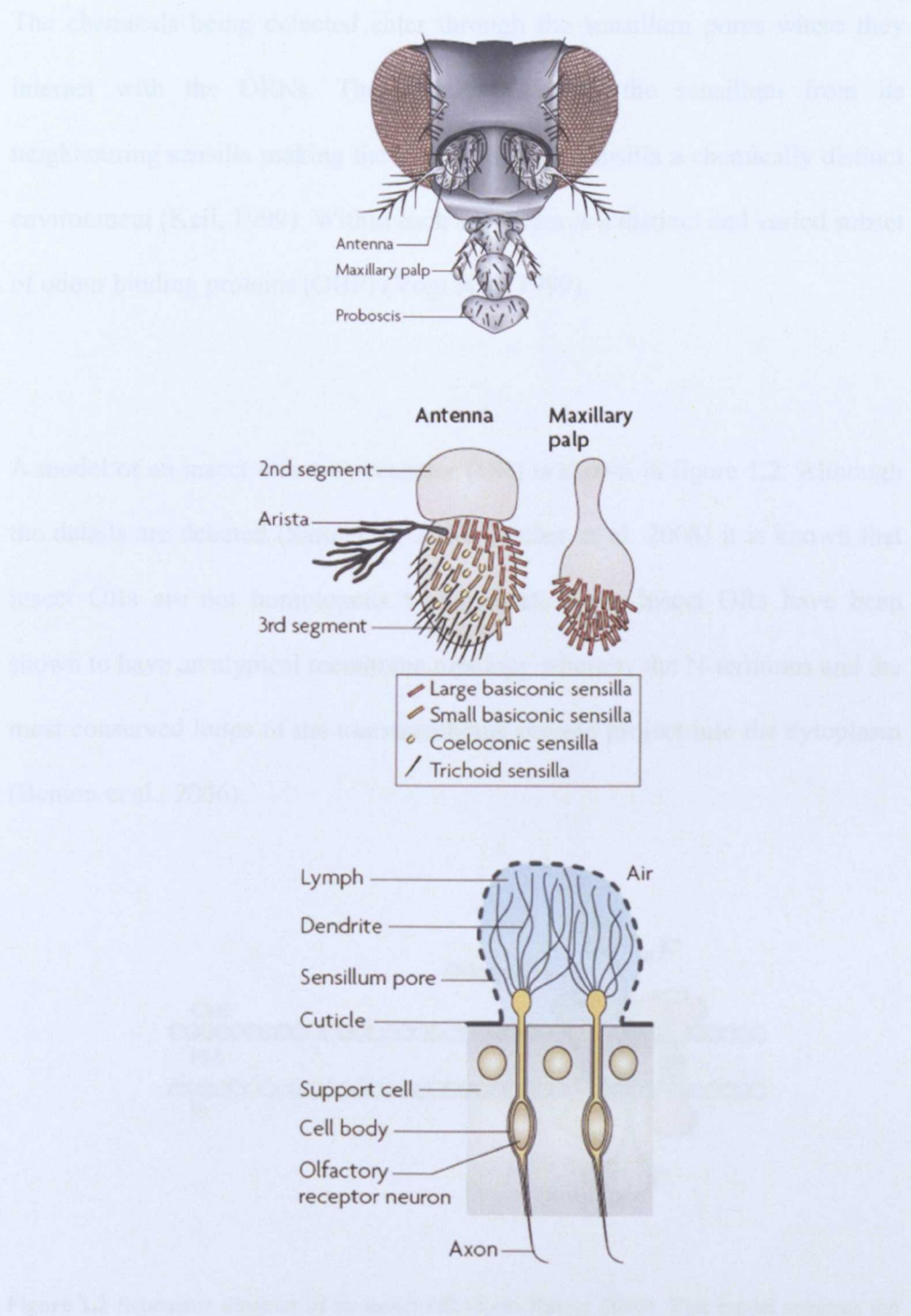


Figure 1.1 Key components of the insect olfactory system. Top: Location of the antennae and maxillary palp which contain the sensilla. Middle: Location of the three types of sensilla. Bottom: Odours enter through the sensillum pore and interact with the olfactory receptor. The cuticle seals each sensillum from its neighbours. Adapted from (Kaupp, 2010).

The chemicals being detected enter through the sensillum pores where they interact with the ORNs. The cuticle seals off the sensillum from its neighbouring sensilla making the interior of each sensilla a chemically distinct environment (Keil, 1999). Within each sensillum is a distinct and varied subset of odour binding proteins (OBP) (Vogt et al. 1999).

A model of an insect olfactory receptor (OR) is shown in figure 1.2. Although the details are debated (Sato et al. 2008; Wicher et al. 2008) it is known that insect ORs are not homologous to vertebrate ORs. Insect ORs have been shown to have an atypical membrane topology whereby the N-terminus and the most conserved loops of the transmembrane protein project into the cytoplasm (Benton et al., 2006).

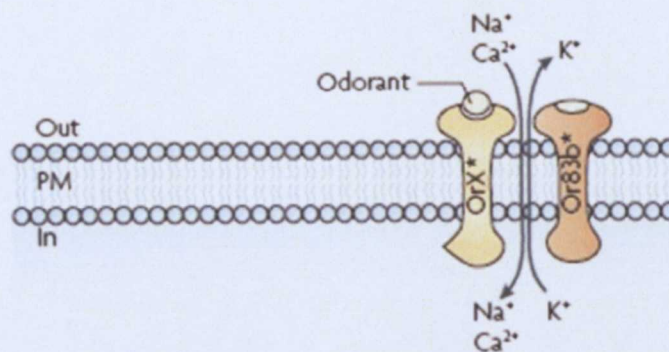


Figure 1.2 Schematic diagram of an insect OR (from Kaupp 2010). This model suggests the OR and co-receptor OR83b form an ion channel that responds to the binding of an odour (Sato et al. 2008).

It is known that the co-receptor OR83b is essential for olfaction and it is thought that the conserved loops of the OR interact with OR83b leading to the protein functioning as a heterodimer.

There are a variety of different olfactory receptors, each detecting particular odours, these lead to a neuronal signal travelling along the ORN which terminates in a distinct area called a glomerulus (labelled GL in figure 1.3) (Hildebrand & G.M. Shepherd, 1997). The colouring in figure 1.3 indicates the specificity involved, for example, the “orange” odour binds to an “orange” ORN, which leads to a glomerulus for all ORNs of this type. The glomeruli are found within the antennal lobe (figure 1.3). From the antennal lobe signals travel to secondary centres in the rest of the brain where signals combine to form the overall “odour image” (Hansson 2002).

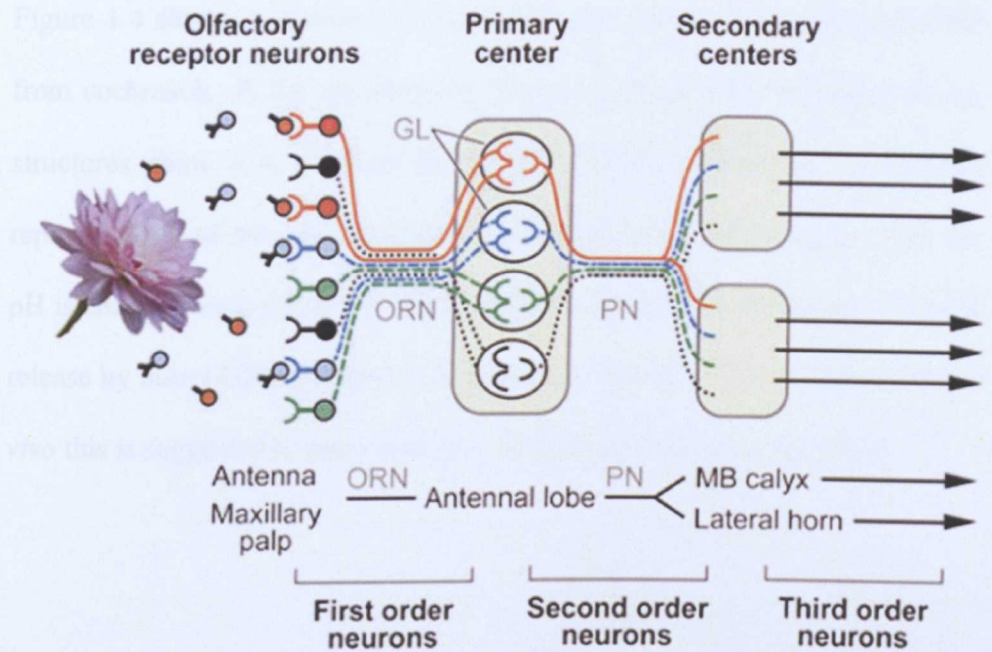


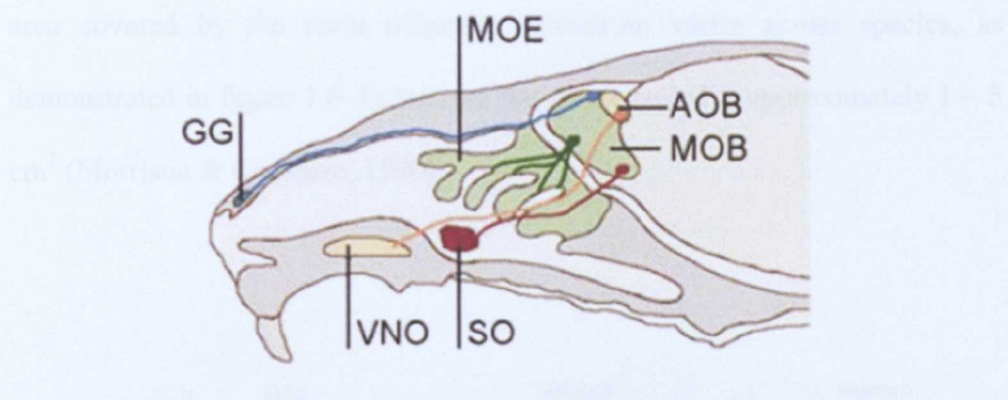
Figure 1.3 Neuronal layout of the insect olfactory system. Receptors are shown as semi circles and odours as filled circles. GL = glomerulus, ORN = olfactory receptor neuron, PN = projection neuron MB = mushroom body (Vosshall & Stocker, 2007).

In insects the role of OBPs is better understood than in vertebrates. Specificity for particular odour and pheromone types has been shown (Vogt et al. 1999). The OBPs of insects were classified on the basis of their sequences and the ligands bound, differentiating Pheromone Binding Proteins (PBPs) from General Odorant Binding Proteins (GOBPs) (Krieger et al. 1996). The first insect OBP was isolated from the giant moth *A. polyphemus* (Vogt & Riddiford 1981) and since then a number of insect PBPs and GOBPs have been identified and structurally characterised (Lartigue et al. 2004; Mohanty et al. 2004; Pesenti et al. 2008; Thode et al. 2008; Zhou et al. 2009; Pesenti et al. 2009; Tsitsanou et al. 2011)

Figure 1.4 shows a number of insect OBPs and PBPs. Aside from LmaPBP from cockroach, all the structures shown are ligand bound. In general all the structures show a 6 α -helical bundle. In the lower panel is a schematic representation of how the structure of *A. polyphemus* PBP changes when the pH is changed from 6.3 to 5.2. This represents a general mechanism for ligand release by insect OBPs, whereby the ligand is released as the pH is lowered (*in vivo* this is suggested to occur near the membrane) (Zubkov et al., 2005).

1.1.3 The vertebrate olfactory system

For vertebrates, the nose and nasal cavity are anatomically important in olfaction. Within this there are a number of important structures which are shown in figure 1.5. Two main pathways are shown; the main olfactory pathway and the accessory olfactory pathway. The feature that separates these two system is the phase in which the detected stimuli is found. The main olfactory system detects gaseous odours, whilst fluid-based stimuli (including pheromones) are detected by the accessory olfactory system (Mucignat-Caretta, 2010). The main olfactory bulb (MOB) is found in the brain and receives nerves impulses from receptors in the main olfactory epithelium (MOE) and the septal organ (SO). Very little is known about the septal organ except that it is a small area of epithelium contain olfactory receptors (Ma et al., 2003). The accessory olfactory bulb (AOB) receives nerve impulses from the vomeronasal organ (VO) (an area dense in pheromone receptors) (Mombaerts 2004) and Grüneberg ganglion (GG) (an area dense in nerve cells, which is important for detecting pheromones that indicate danger) (Brechtbühl et al., 2008) . Each area (excluding the olfactory bulbs) expresses different types of olfactory receptors which are shown in the lower panel of figure 1.5.



	OR	TAAR	V1R	V2R	GC-D
MOE	○	○			○
VNO	○		○	○	
SO	○				
GG	○	○		○	

Figure 1.5 Representative diagram of a cross section through the rodent olfactory system (not to scale). The main olfactory epithelium (MOE) is shown in green. The axons of the MOE and septal organ (SO) are connected to the main olfactory bulb (MOB). The vomeronasal organ (VNO) and Gruneberg ganglion (GG) axons are sent to the accessory olfactory bulb (AOB). The bottom panel indicates where olfactory receptors (OR), trace amine associated receptors (TAAR), vomeronasal receptors (V1R and V2R) and the guanyly cyclase D receptor are found (Kato & Touhara, 2009).

The olfactory epithelium is bathed in the nasal mucosa secreted from the Bowman's gland (which can be seen most clearly in figure 1.7) (Morrison & Costanzo 1990; Lewis & Dahl 2003). In humans the mucus layer is about 60 μM thick (Ohloff, 1994). It is within the nasal mucosa of vertebrates that OBPs are found at high concentrations (100 μM to 1 mM) (Steinbrecht, 1998). The

area covered by the main olfactory epithelium varies across species, as demonstrated in figure 1.6. In humans the area covered is approximately 1 – 5 cm² (Morrison & Costanzo, 1990).

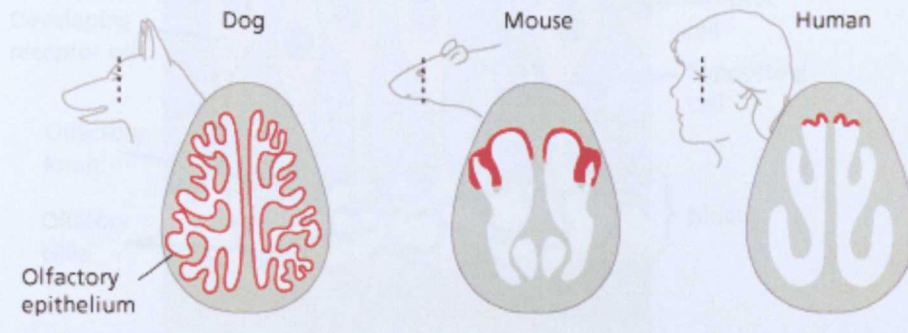


Figure 1.6 Variation in the area of the olfactory epithelium across species.

In contrast to invertebrates, the olfactory receptor neurons (ORNs) are not isolated. Instead they all project into the mucus layer where their cilia “swim” in the olfactory mucus, although they are not motile and do not beat. Mammalian cilia are 30 – 60 μm long and approximately 300 nm in diameter (Jenkins et al., 2009). The location of the ORNs and cilia is shown in the schematic cross section of the olfactory epithelium in figure 1.7.

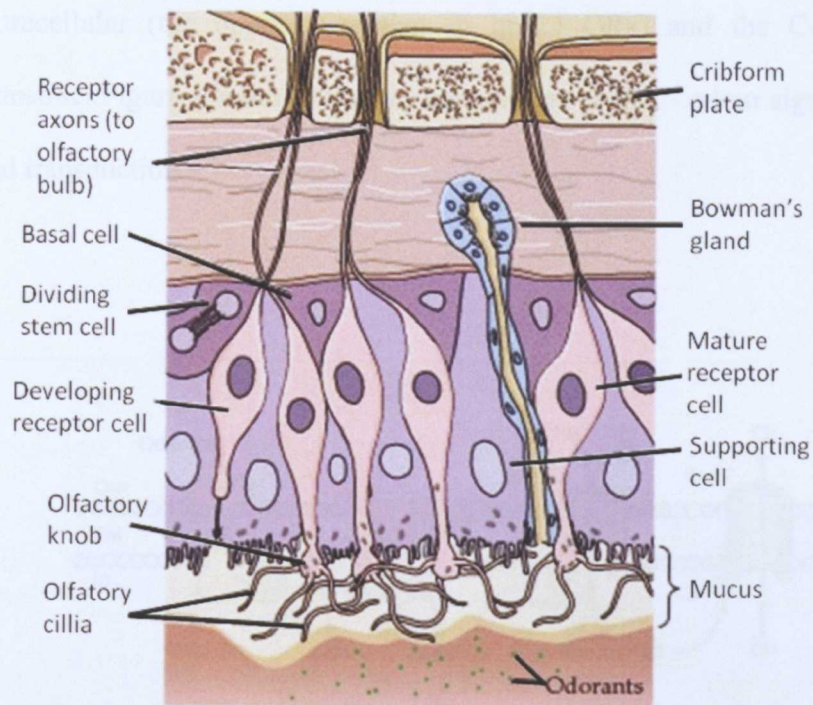


Figure 1.7 Representative diagram of a cross section through the olfactory epithelium (not to scale). The receptor cells have cilia which all project into the mucus layer. The Bowman's gland which secretes mucus is also illustrated. (Purves et al., 2001)

Each of the ORNs expresses one type of olfactory receptor (OR) (Buck et al. 1991). In humans there are over 350 functional OR genes (with a further 400 genes or more designated as pseudogenes) (Malnic et al. 2004), whilst in mice there are 1000 OR genes (Zhang & Firestein, 2002; Young & Trask, 2002; Godfrey et al., 2004).

The ORs of vertebrates are G-protein coupled receptors (GPCRs) (Buck et al. 1991). GPCRs have seven transmembrane α -helices which are connected by intra and extra cellular loops (Grigorieff et al., 1996). The N-terminus is

extracellular (the opposite to that in insect ORs) and the C-terminus is cytosolic. Figure 1.8 shows the key components in OR - odour signal detection and transduction.

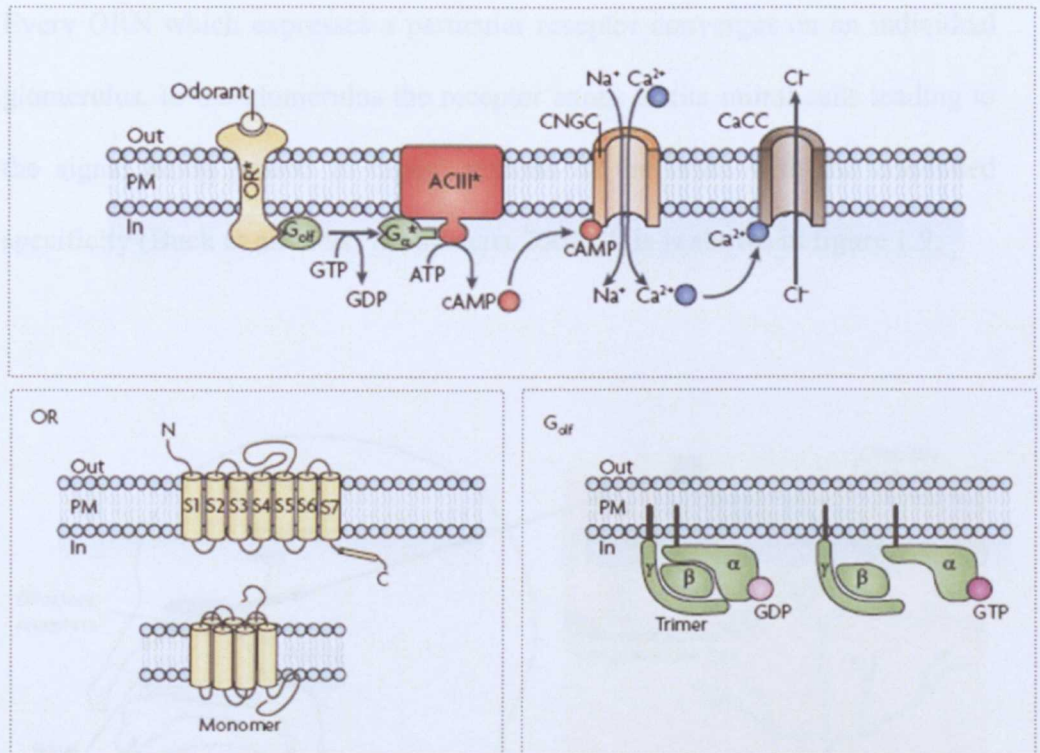


Figure 1.8 Schematic diagram of olfactory receptor (OR) activation in vertebrates by an odour. The structure of the OR is shown diagrammatically in the bottom left panel and shows the seven transmembrane domains of the G-protein coupled receptor. Odour binding to the OR activates G_{olf} (shown in the bottom right panel) which is an olfactory specific G-protein. The α -subunit of G_{olf} dissociates and activates adenylyl cyclase type III (ACIII), which, via cyclic adenosine monophosphate (cAMP), activates the olfactory cyclic nucleotide-gated channel (CNGC) and a Ca^{2+} -activated Cl^- (CaCC) channel. Depolarization of the neuron occurs when both of these channels are activated. Adapted from (Kaupp, 2010).

Studies have shown that odours interact with α -helices 2 – 7 (Katada et al., 2005; Schmiedeberg et al., 2007; Kato et al., 2008) with between 22 and 85 residues predicted to form the binding pocket (Kaupp, 2010).

Every ORN which expresses a particular receptor converges on an individual glomerulus. In the glomerulus the receptor axons excite mitral cells leading to the signal being passed to higher regions of the brain with the continued specificity (Buck et al. 1991; Mombaerts 2004), this is shown in figure 1.9.

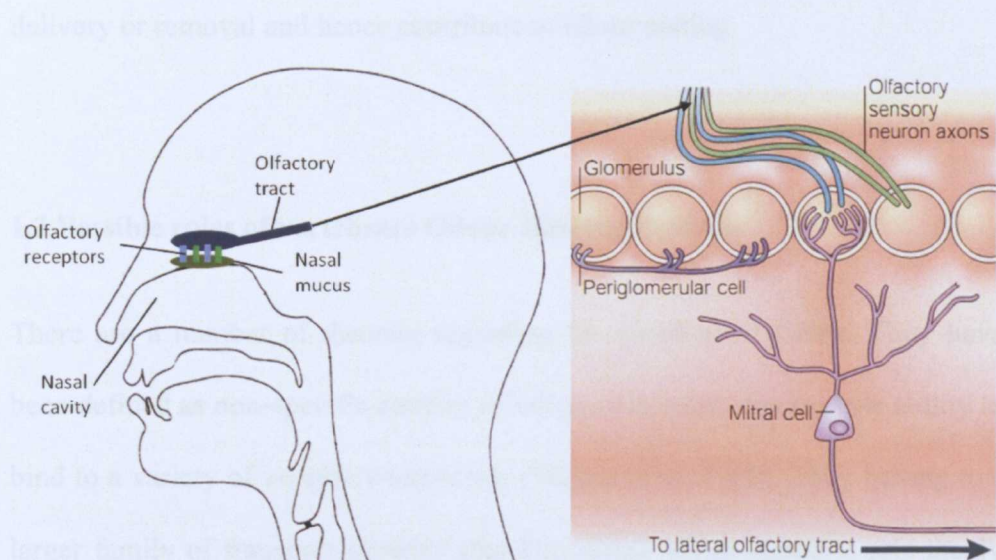


Figure 1.9 Overview of the human olfactory system. The left pane shows a cross section through the nasal cavity whilst the right pane is a cross section through the olfactory bulb (not to scale). The lateral olfactory tract leads to higher areas of the brain (Mombaerts 2004).

How the odour “picture” is built up is still not fully understood. Humans are able to smell up to 10 000 different odour molecules with no two odours

smelling the same (Zarzo, 2007). How this level of distinction occurs is explained by “odour coding”. The principle of odour coding is that each odour activates a particular set of ORs which are used in combination to detect and encode the odour identity (Malnic et al. 1999). The ORs that an individual expresses will of course change the perceived odour. The perception of an odour also depends on how strongly an odour binds and how long the activation lasts for (Bozza & Mombaerts 2001). It is therefore a result of the rate at which odours arrive at the receptors and the rate at which they are removed from the receptor. OBPs are found in the mucus surrounding the receptors so it is possible, like insect OBPs, that they play a part in odour delivery or removal and hence contribute to odour coding.

1.2 Possible roles of vertebrate Odour Binding Proteins

There are a number of theories regarding the function of OBPs. They have been defined as non-specific carriers of odour molecules, due to their ability to bind to a variety of volatile compounds (Tegoni et al. 2000). They belong to a larger family of transport proteins (the lipocalins) which supports this theory (Flower et al. 2000). It is supposed that OBPs help odours cross the mucus layer, however there are a number of hydrophilic odours that will easily dissolve in this layer (Zarzo, 2007). Additionally studies have shown that odours can activate ORs without OBPs being present (Wetzel et al. 1999; Malnic et al. 1999). There are however, hydrophobic odours that are less able to diffuse across the mucus layer.

An important consideration, when determining the exact role of OBPs, is whether or not they bind to ORs, and if binding depends on the OBP being pre-bound to an odour. It has been suggested that, under physiological conditions, unliganded OBP (from pig) can selectively bind a human OR (Matarazzo et al., 2002). More recently Vidic et al. (2008) used yeast-derived nanosomes grafted onto a sensorchip and surface plasmon resonance (SPR) to demonstrate OBP – OR interactions, as well as the release of the OBP upon specific odour binding to the OR. This idea would make the role of OBPs more complex, and certainly make these proteins important in odour coding, as they would be directly regulating the ORs. Evidence is not extensive and as yet does not suggest if the possible OR-OBP interaction is selective.

Whether or not the role of OBPs involves the binding to ORs, the concept of OBPs being “odour sinks” is still relevant. This idea involves OBPs increasing the concentration of odours when they are present at initially low concentrations; whilst free odours at initially high concentrations are able to travel unassisted across the mucus layer (Zarzo, 2007). It has also been suggested that OBPs may play a role in preventing receptor saturation at high odorant concentration (Matarazzo et al., 2002). Also key is the kinetics of the release of the odour by the OBP, which in turn may regulate the rate of odour uptake by the ORs.

The theory of OBPs playing a role as odour “scavengers” has also been reported (Grolli et al. 2006), whereby OBPs are suggested to bind high levels

of toxic odours so they can be removed from the nasal cavity, preventing an over-sensitization of the ORs. However there are already a number of different types of enzymes identified as being present in the olfactory epithelium that degrade and biotransform volatile odour compounds once they have been perceived (Breer 2003).

The presence of multiple OBP subtypes identified in some species raises the possibility of OBPs being selective filters involved in the pre-selection of odours. In the porcupine *Hystrix cristata* eight OBPs were identified (Felicoli et al., 1993). In rat three subtypes have been identified (Löbel et al. 1998; Löbel et al. 2001) and studies have identified each subtype as having different odour preferences (although some are overlapping) (Löbel et al. 1998; Löbel et al. 2002). Similar odour preference (rather than broad binding) was also shown for the human OBP-2A (Tcatchoff et al., 2006). In terms of odour coding the idea of pre-sorting of odours adds another stage and hence more complexity. Rather than just the activation of the ORs being important, there would be a temporal effect due to different binding rates of different OBP subtypes. It is important to consider, however, that despite there being more than 350 ORs in humans; only two subtypes of OBPs have been identified. As yet no correlation has been found between the repertoire of odours an organism can distinguish and the number of OBPs it has.

Major urinary proteins (MUPs) have received detailed study (Lücke et al. 1999; Kuser et al. 2001; Phelan et al. 2010; Perez-Miller et al. 2010) and owing to the high degree of similarity observed between MUPs and some OBPs

(Spinelli et al. 1998) their role should be considered. MUPs are known to bind to pheromonal compounds (Bocskei et al., 1992; Židek et al., 1999; Timm et al., 2001). One of the roles MUPs are known to play is in the sexual attraction of female mice by male mice. Pheromone bound MUP is secreted into the urine of a male mouse, when the pheromone is released from the MUP it is available to be detected by the vomeronasal organ of a female mouse, triggering a behavioural response and resulting in sexual attraction (Flower 1996).

In this case the tight binding of the pheromone to MUP is essential as it allows the signal to persist for long enough to be detected and elicit a response. A “nasal” MUP has recently been identified in the olfactory system of rodents (Perez-Miller et al., 2010) and this may indicate a further role in pheromonal receptor delivery or removal from the receptor, in a similar manner to that suggested for OBPs.

1.3 The structure of vertebrate Odour Binding Proteins

1.3.1 Lipocalins

The structures of vertebrate OBPs (and MUPs) are classified as being part of the lipocalin superfamily. Lipocalins as a protein family show very low sequence similarity (frequently less than 20%), however they possess a conserved structure in the form of a six or eight stranded anti-parallel β -barrel, and are functionally similar in their ability to bind small hydrophobic, or lipophilic, molecules (Flower et al. 2000)

The family belongs to a larger superfamily, the calycins, which groups together proteins that contain a β -barrel which binds small hydrophobic ligands. The superfamily also includes fatty acid-binding proteins (FABPs), the avidins, a group of metalloprotease inhibitors (MPIs) and triabin. (Flower 1996; Flower et al. 2000; Fuentes-Prior et al. 1997).

Lipocalins are typically 18-20 kDa in size, with some larger lipocalins possessing additional structural elements and domains that confer specific functions. The general lipocalin structure (shown in figure 1.10) is that of a calyx, formed by the folding of a β -sheet consisting of eight anti-parallel strands, in a (+1) topology, (figure 1.10) labelled A to H, which hydrogen bond to form a barrel. There is an “open” end to the internal ligand binding site. This is closed on the other side by the N-terminal of the molecule (which contains a 3^{10} helix) crossing to strand A.

Other key features include a larger α -helix present on the C-terminal side of strand H. The loops connecting the β -strands are all of the hairpin type, aside from loop 1 (connecting strands A and B), which is a more flexible Ω loop, and is suggested to form a “cover” or “lid” (Grzyb & Latowski, 2006)

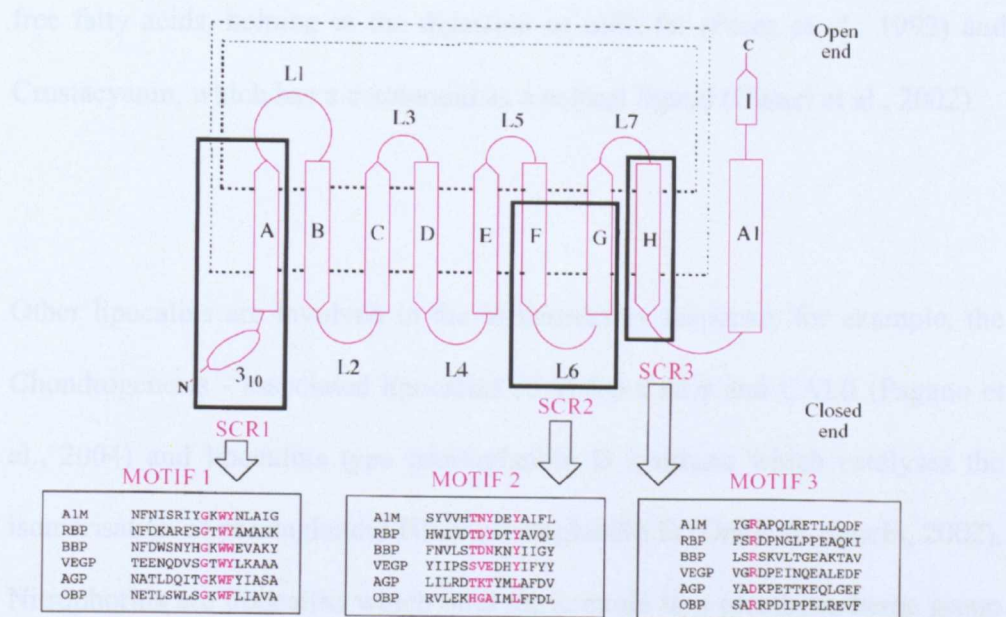


Figure 1.10 General secondary structure of a lipocalins highlighting structurally conserved regions (SCRs). Eight anti-parallel strands form a β -barrel (labelled A – H) with a C-terminal α -helix (A1) and a 3^{10} helix at the N-terminus. In some lipocalins an additional β -sheet is identified (labelled I). The loops connecting the β -strands are labelled L1-L6. L1 is an omega loop, which is more flexible. Kernel lipocalins (A1M = α_1 -microglobulin, RBP = retinol binding protein and BBP = bilin binding protein) and outlier lipocalins (VEGP = von Ebner's gland protein, AGP = α_1 -acid glycoprotein and OBP = odour binding protein (bovine)) are shown. (Flower 1996)

As well as being sequentially divergent, members of the lipocalins family also diverge in the roles they play. Their physiological significance is not only as proteins that transfer lipophilic molecules (Grzyb & Latowski, 2006), though this description itself includes a broad range of individual roles. Examples include; Retinol Binding Protein (Newcomer et al., 1984; Zanotti et al., 1993; Redondo et al., 2008), β -lactoglobulin, which is postulated to bind and remove

free fatty acids, helping in the digestion of milk fat (Perez et al., 1992) and Crustacyanin, which has a carotenoid as a natural ligand (Cianci et al., 2002).

Other lipocalins are involved in the inflammatory response; for example, the Chondrogenesis - associated lipocalins (CALs) CAL γ and CAL β (Pagano et al., 2004) and lipocalins type prostaglandin D synthase which catalyses the isomerisation of prostaglandin H₂ to prostaglandin D (Urade & Eguchi, 2002). Nitrophorins are lipocalins which bind nitric oxide to a prosthetic heme group located at the “open” end of their β -barrels (Andersen et al., 1998). Histamine binding proteins found in the saliva of hematophagus (blood-feeding) parasites are identified as lipocalins despite histamine being a hydrophilic ligand (Sangamnatdej et al., 2002).

Lipocalins are separated into two subfamilies; the closely related kernel lipocalins and the outlier lipocalins which are more divergent. From the examples above Crustacyanin, Nitrophorins and Histamine binding proteins are defined as outlier lipocalins. Odour binding proteins are also outlier lipocalins (Flower et al. 2000; Grzyb & Latowski 2006). The main division between these groups are the SCRs (structurally conserved regions) rather than their physiological roles. In figure 1.10 the sequence motifs which correspond to the SCRs are shown. The top 3 sequences (A1M, RBP and BBP) are kernel lipocalins and the bottom 3 sequences (VEGP, AGP and OBP) are outlier

lipocalins. Generally outlier lipocalins only possess two of the three SCRs, and it is often motif 2 which is different.

The vertebrate OBPs all share a short -G – X – W- motif located near the N-terminus (generally 15 to 20 residues away), however sequence identity overall is usually below 50% (Pelosi 1994; Tegoni et al. 2000). An alignment of the primary sequences of the most studied OBPs is shown in figure 1.11. As well as the G-X-W motif, all the OBP share a conserved lysine residue around position 72. There is also a Y-X-X-X-Y-X-G motif that is conserved with the exception of the rat OBP2 and Human 2A proteins. With the exception of bovine OBP which has no cysteines, all of the OBPs have a pair of conserved cysteines (one close to the C-terminus and the other near position 63).

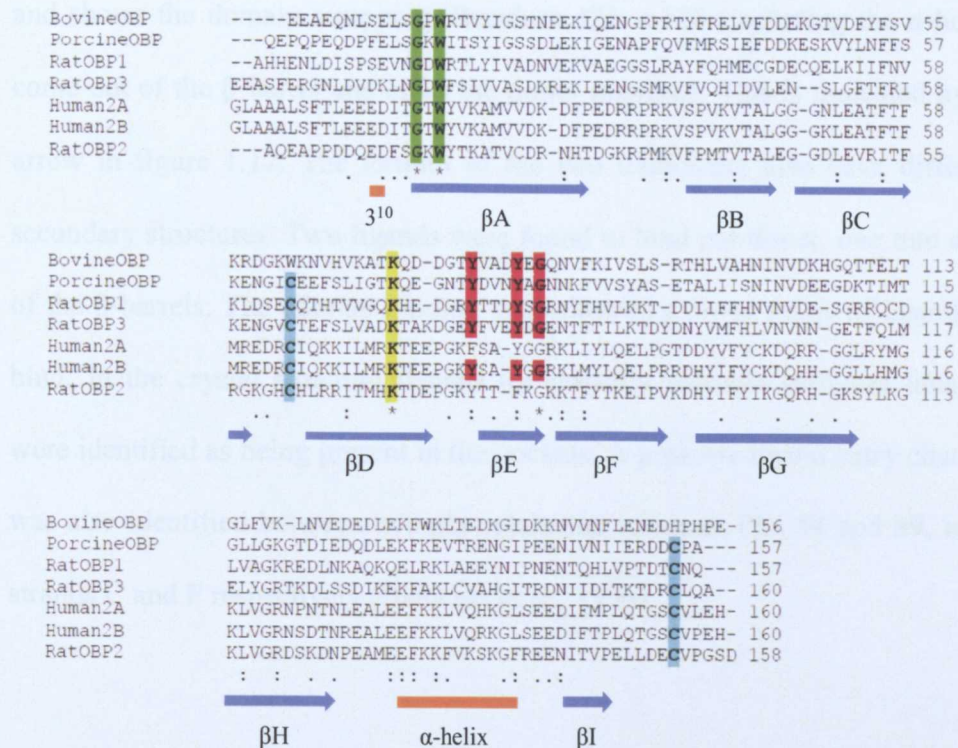


Figure 1.11 Alignment of the most well studied vertebrate OBPs. * = identical residue, : = equivalent residue and . = similar type of residue. Conserved residues are highlighted and emboldened. The G-X-W motif is in green. The conserved cysteines are highlighted in blue, whilst the conserved lysine is shown in yellow. The Y-x-x-x-Y-x-G motif is shown in red.

1.3.2 Bovine and Porcine OBP

Bovine OBP was the first vertebrate OBP identified and was found in cow nasal mucosa through its binding to 2-isobutyl-3-methoxypyrazine (Bignetti et al. 1985). The structure of Bovine OBP (bOBP) was also the first vertebrate OBP structure to be solved, and the structure showed that bOBP is a domain swapped dimer, with an approximate two-fold axis of symmetry. The unusual lack of cysteine residues is most likely a requirement for the domain swapping to occur (Tegoni et al. 1996). The structure of bOBP is shown in figure 1.12,

and shows the domain swapping. Residues 123 – 126 (including the α -helix) come out of the β -barrel and cross the dimer interface. This is indicated by an arrow in figure 1.12. The termini of the two monomers also have different secondary structures. Two ligands were found to bind per dimer, one into each of the β -barrels. The odorant binding sites allowed a broad range of ligands to bind. In the crystal structure ligands (or possibly multiple different ligands) were identified as being present in the pockets. A possible ligand entry channel was also identified between two phenylalanine residues, Phe 54 and 89, in β -strands C and F respectively (Bianchet et al., 1996).

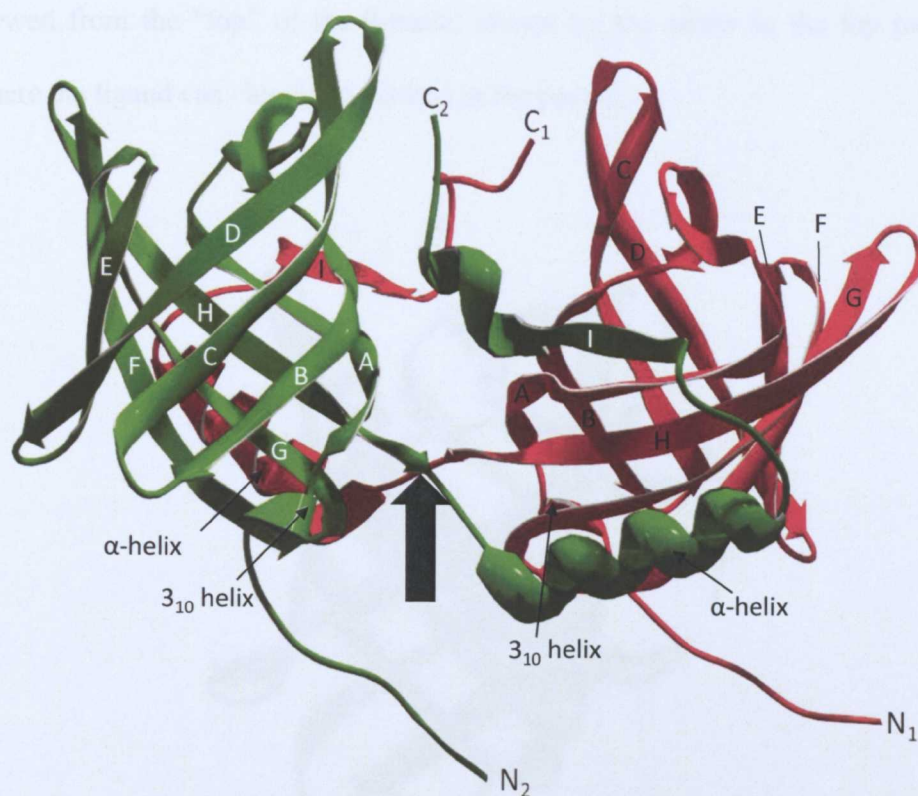


Figure 1.12 Crystal structure ribbon diagram of Bovine OBP. One monomer is shown in green (with white β -strand labels), whilst the other is in red (with black β -sheet labels). The domain swapping is highlighted by the arrow. Figure constructed from the 1OBP Protein Databank (PDB) file.

Crystal structures of bOBP in complex with a number of odorant molecules have been solved (Vincent et al. 2004) and used to contrast bOBP with porcine OBP (pOBP), which is monomeric and unlike bOBP did not purify with a ligand already bound. Two monomers of pOBP were found in each unit cell (Spinelli et al. 1998; Vincent et al. 2000). The crystal structure of pOBP bound to 2-isobutyl-3-methoxy pyrazine (IBMP) (Vincent et al. 2000) is shown in figure 1.13. The upper panel shows the β -barrel, and α -helix behind, demonstrating the typical lipocalin fold. The bottom panel shows the pocket

viewed from the “top” of the β -barrel shown by the arrow in the top panel where the ligand can clearly be viewed in the pocket.

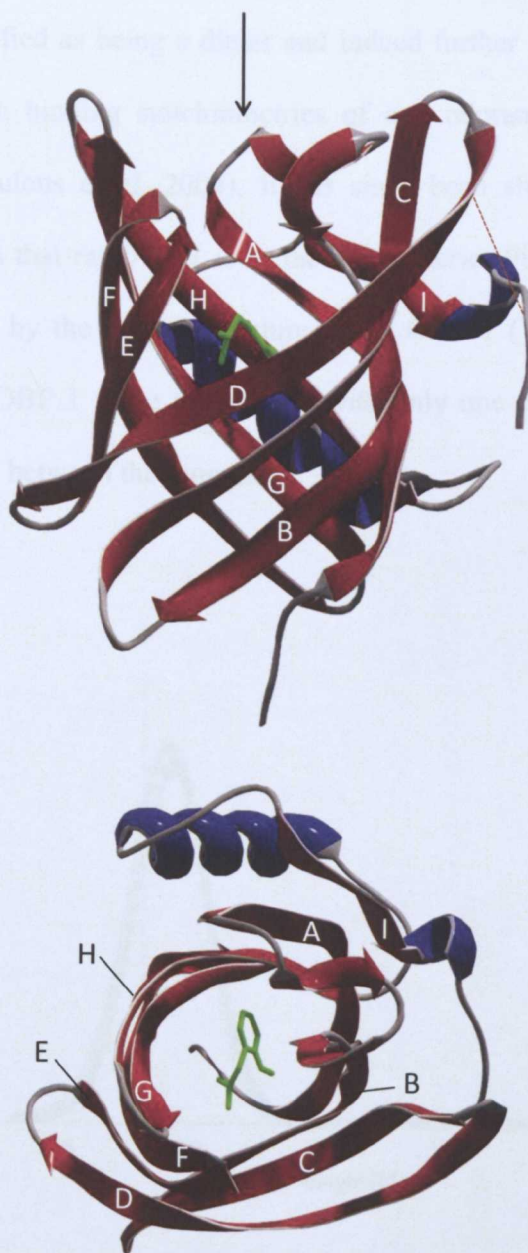


Figure 1.13 Ribbon diagram of the crystal structure of Porcine OBP. Constructed from the 1DZK PDB file. The β -strands are labelled (A-I) and helical content is shown in blue. The disulphide bond tethering the C-terminal to β -strand D is shown by a dotted line. The ligand 2-isobutyl-3-methoxypyrazine is shown in green. The bottom panel shows the pocket viewed from above (denoted by the arrow in the upper panel).

1.3.3 Rat OBP 1

As well as in cow and pig, OBPs have also been identified in rat. Initially rat OBP 1 was identified as being a dimer and indeed further studies were based on this idea, with binding stoichiometries of one odorant per dimer being identified (Nespoulous et al. 2004). It has since been shown by analytical ultracentrifugation that rat OBP 1 is in fact monomeric (Figure 1.14). This is further confirmed by the crystal structure of rat OBP 1 (White et al., 2009) which shows rat OBP 1 to be monomeric with only one possible salt bridge being able to form between the monomers.

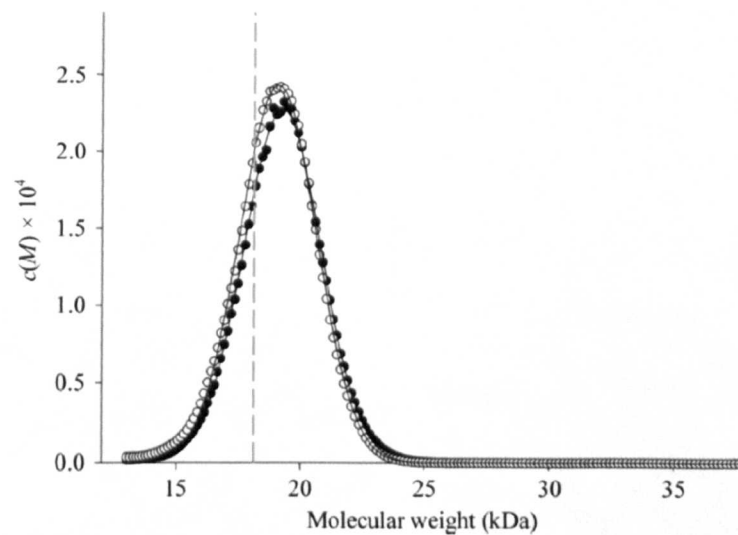


Figure 1.14 Analytical ultracentrifugation used to demonstrate that rat OBP1 is a monomer. The apo form is shown in open circles and the protein saturated with a two-fold molar excess of linalool is shown in closed circles. The dashed line indicates the calculated molecular weight of an OBP1 monomer. Figure from White et al. (2009).

The crystal structure of rat OBP1 is shown in figure 1.15. Rat OBP1, like bovine OBP was crystallised with an electron density (attributed to a yeast pigment from the purification procedure) already in the pocket. Rat OBP 1, like pOBP, is a classical lipocalin. It has an eight stranded anti-parallel β -barrel with an N-terminal 3^{10} helix covering the bottom opening of the barrel, and an α -helix at the C-terminus, leaving a possible opening at the “top” of the barrel. However OBP1, in contrast to the other sequences aligned in figure 1.13, has five cysteine residues. Two disulphide bridges are formed (shown as yellow sticks in figure 1.15) and leaving one free cysteine. It was found that the majority of residues in the binding pocket of rat OBP 1 were aromatic; this is consistent with porcine and bovine OBP. To date bOBP, pOBP and rat OBP1 are the only vertebrate OBPs for which the structures have been solved.

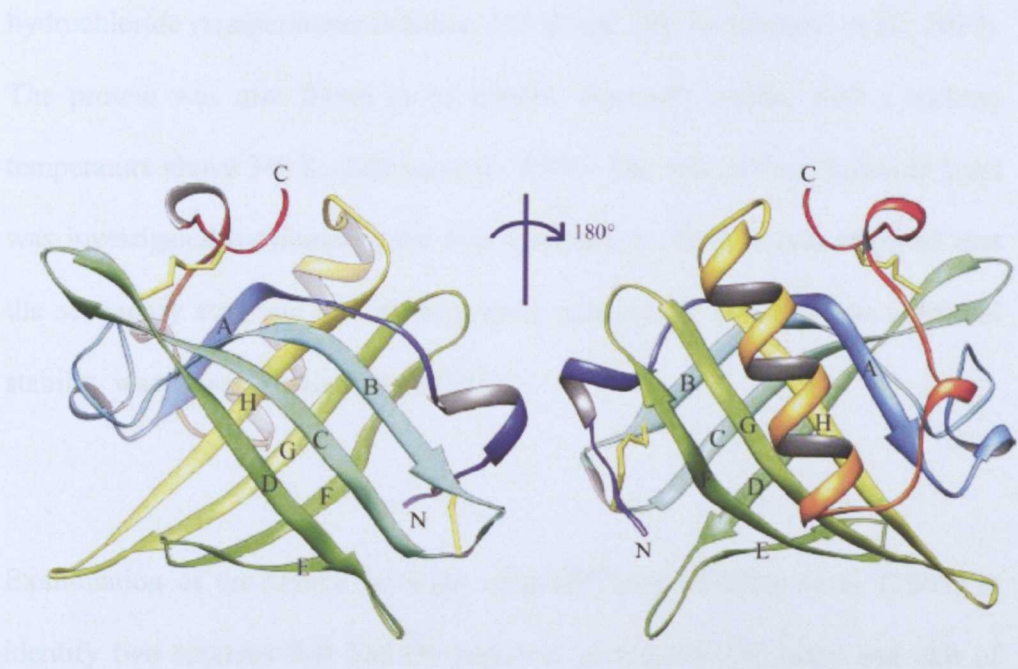


Figure 1.15 Crystal structure of OBP1. The β -sheets are labelled A-H. The α -helix is shown in black and orange. Yellow sticks denote the disulphide bridges. Figure from White et al. (2009).

1.3 Previous studies of Odour Binding Proteins

A number of previous studies of vertebrate OBPs have focussed on porcine OBP (pOBP), particularly, as it is a more representative of OBPs as lipocalins than bovine (bOBP). As well as making comparisons between pOBP and bOBP (Spinelli et al. 1998), other studies have looked at a range of biophysical characteristics of pOBP. The phosphorescence of the tryptophan residue (part of the G-X-W motif) was examined and used to conclude that pOBP exists in a range of conformations that slowly interconvert (D'Auria et al., 2008). The stability of the protein was assessed using a chemical denaturant and it was found that pOBP unfolds at a concentration of between 1.4 – 3.5 M guanidine hydrochloride (temperatures between 293 K and 298 K) (Staiano et al., 2007). The protein was also found to be relative thermally stable, with a melting temperature above 340 K (Burova et al., 1999). The role of the disulphide bond was investigated by mutating the two cysteines to alanines and revealed that the secondary structure and binding were maintained, however the chemical stability was lower (Parisi et al., 2005).

Examination of the crystal structure of pOBP lead Meillour et al. (2009) to identify two residues that had the potential to regulate the entry and exit of ligands into the binding pocket. The results of molecular dynamics simulations looking at the binding of undecanal are shown in figure 1.16. They revealed that phenylalanine 35 appeared to reorientate undecanal towards the exit whilst tyrosine 82 allowed complete release of undecanal from the pocket. Whilst

tyrosine 82 is part of the conserved Y-X-X-X-Y-X-G motif in OBPs, phenylalanine 35 is only conserved between pOBP and bOBP. Mutants were also constructed (Y82A, F35A and Y82A/F35A) however the fluorescent probe 1-aminoanthracene (1-AMA) was not taken up, so displacement studies could not be conducted. This result indicated that the mutations had a serious effect on binding.

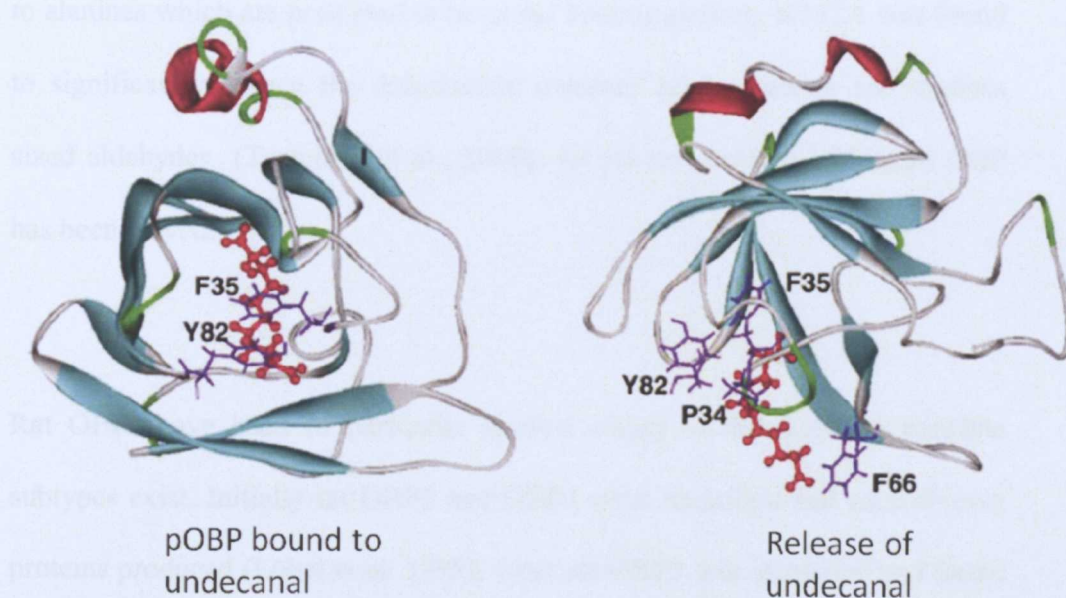


Figure 1.16 Molecular dynamics simulations of undecanal unbinding from porcine OBP. On the left the residues Y82 and F35 cover the pocket opening. These residues move in order to release the ligand (right).

In other studies it has been suggested through molecular dynamic simulations that pOBP and rat OBP 1 undergo strand pair opening in the presence of a ligand and it is suggested that this may be required for recognition of OBPs by olfactory receptors (Hajjar et al., 2006). It is yet to be determined

experimentally how a ligand enters the binding pocket using this model and how this is regulated, as well as how the ligand is released again and how this is controlled.

OBPs have also been identified in human nasal mucosa (Briand et al. 2002) and hOBP-2A has been shown to have specificity for aldehydes, conferred by a lysine residue in the binding pocket. This was studied by mutating three lysines to alanines which are predicted to be in the binding pocket,. K112A was found to significantly reduce the dissociation constant of the protein for medium sized aldehydes (Tcatchoff et al., 2006). As yet no structure of human OBP has been solved.

Rat OBPs have been of particular interest owing to the fact that multiple subtypes exist. Initially rat OBP1 and OBP2 were identified and recombinant proteins produced (Löbel et al. 1998). Later rat OBP3 was identified and found to bind well to heterocyclic compounds (Löbel et al. 2001). Within the rat subtypes some degree of ligand binding specificity has been identified (Löbel et al. 1998; Löbel et al. 2001). It has been suggested that OBP1 may bind to odours that are aromatic hydrocarbons, OBP2 may have a preference for aliphatics and nitriles, whilst somewhat broadly, OBP3 binds well to compounds containing a ring structure (Löbel et al. 2001). This of course would indicate that OBPs may be selectively “pre-sorting” odours that are delivered to the ORs and would make them an important part of fully understanding olfaction. This is particularly interesting as OBP1 and OBP3

were found to be expressed in the same location (the incisor root) (Löbel et al. 2001), suggesting from an evolutionary point of view they are likely to perform different roles.

Of the three rat subtypes, OBP1 has been studied in the most detail. The crystal structure of OBP1 revealed that tyrosine 82 was positioned to “cap” the entrance to the ligand binding pocket (White et al., 2009) in a similar manner to that suggested for pOBP, and found previously by molecular dynamics simulations of OBP1 (Hajjar et al., 2006; Golebiowski et al., 2006).

Rat OBP1 has also been used to study the kinetics of odour uptake by the protein, and the liquid to air transfer rates of odours. Binding and release of odours was monitored using a cellulose reaction surface onto which OBP1 was placed, housed within a reaction vessel and connected to an atmospheric pressure chemical ionization mass spectrometer (APCI-MS). The mass spectrometer detected the odorants that were released. This study proposed that a rapid off-rate regime (of odour release by the protein) is required *in vivo* in order correlate with the rates involved in the olfactory system. (Borysik et al., 2010).

The initial identification of the third rat subtype (OBP3) which is the focus of this study recognised the protein as being similar to the major urinary proteins (MUPs) (Löbel et al. 2001) mentioned in section 1.2, with OBP3 and MUP I

sharing 83% sequence identity. The MUPs are also part of the lipocalin family and have received detailed study.

1.4 Major urinary proteins

Four types of major urinary proteins are known to be found in mice. MUP I is found in the urine, after initially being expressed in the liver and filtered by the kidneys (Knopf et al. 1983; Kuhn et al. 1984; Clark et al. 1984). It binds to pheromonal compounds that are able to induce sexual maturation in mice (Novotny & Wiesler, 1999). MUP II represents a set of pseudogenes and MUP III proteins are suggested to be glycosylated (Bishop et al. 1982; Al-Shawi et al. 1989). MUP IV is expressed in the vomeronasal mucosa (Shahan et al., 1987; Al-Shawi et al., 1989) and, due to this, is dubbed “nasal MUP”.

The structure of wild type MUP 1 was solved by x-ray crystallography (Bocskei et al., 1992) whilst the structure of a recombinant MUP 1 was solved both by NMR (Lücke et al. 1999) and x-ray crystallography (Kuser et al., 2001). The structure of MUP I with ligands bound has also been solved by both NMR and X-ray crystallography (Zidek et al. 1999; Bingham et al. 2004; Homans 2005; Barratt et al. 2005; Barratt et al. 2006). Recently the structures of MUP IV bound to three mouse pheromones were also solved by x-ray crystallography (Perez-Miller et al., 2010).

It is particularly noteworthy to mention that the crystal structure of MUP I contained no ordered solvent molecules. Molecular dynamics simulations also revealed a net expulsion of water placed into the binding pocket (Barratt et al., 2005). Thermodynamic experiments found, surprisingly, that, despite the hydrophobicity of the residues lining the binding pocket and the hydrophobicity of the ligands, binding to MUP I was exothermic, strongly enthalpy dominated and had an unfavourable entropy term (Bingham et al. 2004; Barratt et al. 2005).

Experiments looking at ligand binding to MUPs have focussed on pheromonal ligands, and as such no large scale screens of the possible ligands that could be bound have been undertaken.

Studies on MUPs should be given careful consideration due to the high degree of sequence identity between OBP3 and the MUPs. MUPs are amongst the first proteins for which these unusual interactions, which are enthalpy driven and appear to take place without the net expulsion of water being a requirement, have been found. Detailed, long timescale molecular dynamics simulations have recently been undertaken in an attempt to understand the unusual thermodynamics of MUP. These studies revealed surface loops of the protein changed their flexibility upon ligand binding (some becoming more flexible, others less flexible). In addition a very low solvent density in the absence of a ligand was shown. Once a ligand was bound, it retained a high amount of rotational freedom. These findings are suggested to explain the entropy values

found in previous studies and challenge the previously cited dogma “nature abhors a vacuum” (Roy & Laughton, 2010).

1.5 Aims and objectives of this study

In order to learn more about the binding interactions of Odour Binding Proteins the binding profile of rat OBP3 was investigated using defined sets and series of odours. These experiments was used to establish the thermodynamic characteristics of OBP3-odour interactions and describe the affinity of OBP3 for a number of odours

The 3D structure of OBP3 was modelled in order to relate the thermodynamic characteristics of odour binding to the structure of OBP3. NMR spectroscopy was used to provide information on the binding interactions on a per residue basis. In order to do this the NMR spectrum of OBP3 needed to be assigned. In addition the mutagenesis of several individual residues was used to investigate the possible causes of apparent rat OBP subtype – odour subset preference. The importance of the aromatic nature of the binding pocket of OBPs will also be probed using mutagenesis studies.

2. MATERIALS AND METHODS

2.1 Molecular Biology

All chemicals were obtained from Sigma Aldrich (Poole, Dorset, UK) unless otherwise stated.

2.1.1 Strains

In order to produce OBP3 *Escherichia coli* (*E. coli*) XL1-blue and *E. coli* Top 10 cells (Invitrogen, Paisley, UK) were used for transformation by electroporation and subsequent propagation of plasmid DNA. For transformation of mutagenesis products, XL-10 gold ultracompetent *E. coli* cells were used (Agilent Technologies, West Lothian, UK). For expression of recombinant OBP3, commercially available *E. coli* M15 cells (Qiagen, Crawley, UK) were used. The genotypes of the *E. coli* strains used are shown in table 2.1.

Name	Genotype	Original Source
<i>E. coli</i> XL1-Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac</i> [F' <i>proAB lacIqZΔM15 Tn10</i> (Tetr)]	Invitrogen
<i>E. coli</i> Top-10	F'[<i>lacI^f Tn10(tet^R)</i>] <i>mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 deoR nupG recA1 araD139 Δ(ara-leu)7697 galU galK rpsL(Str^R) endA1 λ⁻</i>	Invitrogen
<i>E. coli</i> XL-10 Gold	<i>endA1 glnV44 recA1 thi-1 gyrA96 relA1 lac Hte Δ(mcrA)183Δ(mcrCB-hsdSMR-mrr)173 tet^R F'[proAB lacI^fZΔM15 Tn10(Tet^R Amy Cm^R)]</i>	Agilent Technologies
<i>E. coli</i> M15	<i>Nals⁺, Strs⁺, Rifs⁺, Thi⁻, Lac⁻, Ara⁺, Gal⁺, Mtl⁻, RecA⁺, Uvr⁺, Lon⁺</i>	L.Briand, INRA/Qiagen

Table 2.1 Strains used in this study

2.1.2 Standard growth conditions and antibiotic selection for OBP3 production

All *E. coli* growth was performed in Luria-Bertani (LB) medium (10 g/l bacto-tryptone, 5 g/l bacto-yeast extract, 10 g/l NaCl) at 37 °C with 200 rpm agitation (in a Gallenkamp, orbi-safe incubator, Loughborough, UK) unless otherwise stated. Ampicillin (100 µg/ml) and kanamycin (50 µg/ml) were added to select for cells containing the plasmids pQE31 and pREP4 respectively. LB agar was produced by adding 15 g/l bacteriological agar to LB broth.

2.1.3 Strain storage

In order to store propagation and expression strains of *E. coli*, glycerol stocks were prepared. To do this, a single colony of the desired cells was picked and grown overnight in 5 ml of LB. The resultant culture was centrifuged at 3 000 g in a bench top centrifuge (Heraeus® Biofuge® Stratos, Thermo Scientific, Wilmington, USA). The supernatant was discarded and the pellet resuspended in a total volume of 2 ml of LB (50 % v/v) and glycerol (50 % v/v) which had been filter sterilized (0.2 µm Minisart® syringe filter, Sartorius, France) This was divided into 50 µl aliquots, snap frozen in liquid nitrogen and stored at -80 °C.

2.1.4 Production of electrocompetent XL1-Blue *E.coli*

Electrocompetent *E. coli* were prepared by the method described in (Sambrook & Russell (2001). Briefly, four 25 ml overnight cultures of XL1-Blue *E.coli*

were used to inoculate four flasks each containing 250 ml of fresh LB, these were grown to an OD_{600nm} of 0.4, after which time the flasks were transferred to an ice water bath for 30 minutes in order to stop growth. The cultures were then transferred to ice-cold centrifuge tubes and from this point on all steps were carried out at 4°C or lower. The cells were harvested at 2500 rpm for 15 minutes before being resuspended in 250 ml sterile H₂O. The cells were then harvested at 2500 rpm for 20 minutes and resuspended in 125 ml of 10 % glycerol, the centrifugation was repeated and the cells resuspended in 10 ml of 10 % glycerol. After a final centrifugation using the same settings the pellet was resuspended in 1 ml GYT (10 % glycerol, 0.125 % (w/v) yeast extract, and 0.25 % (w/v) tryptone). The electrocompetent cells obtained were snap frozen in liquid nitrogen and stored in 40 µl aliquots (at a concentration of 2×10^{10} to 3×10^{16} cells) at -80 °C until required.

2.1.5 Transformation of electrocompetent XL1-Blue *E.coli* cells.

1 µl of plasmid DNA at a concentration of 1 ng/µl was added to 40 µl of electrocompetent cells (on ice) before being pipetted into an ice-cold electroporation cuvette and electroporated with a pulse of 25 µF capacitance, 2.5 kV and 200 ohm resistance (Electroporator 2510, Eppendorf, Cambridge, UK). 1 ml of LB medium was then added and the cells allowed to recover for 1 hour at 37 °C with 200 rpm shaking. The cells were spread on LB agar plates with the relevant antibiotic selection.

2.1.6 Transformation of chemically competent M15 *E. coli* by heat shock

1 µl of plasmid DNA at a concentration of 1 ng/µl was added to 40 µl of chemically competent *E. coli* cells on ice, in 1.5 ml microtubes, and incubated on ice for 30 min before being placed in a 42 °C waterbath for 45 seconds, after which time they were returned to ice for 2 minutes. Recovery was allowed in 1 ml of LB medium for 1 hour at 37 °C with 200 rpm shaking. The cells were spread on LB agar plates containing the relevant antibiotic selection.

2.1.7 Preparation of purified pQE31 plasmid

Plasmid preparation was performed using a QIAprep[®] Miniprep Kit (Qiagen). 5ml cultures of XL1-Blue or XL-10 gold cells containing pQE31::obp3 were grown overnight in LB containing 100 µg/ml ampicillin. These were grown from single colonies. In order to pellet the cells the cultures were centrifuged at maximum speed (13,000 g) in a microcentrifuge (Eppendorf). The supernatant was discarded and the cells resuspended in 250 µl of resuspension solution containing RNase A, complete resuspension was guaranteed by the use of a MixMate[™] tube shaker (Eppendorf). 250 µl of lysis solution was then added and the reaction allowed to proceed for up to 5 min at room temperature. The lysate was then neutralised and cell debris pelleted at 13,000 g. The supernatant was transferred to a spin column assembly again at maximum speed, for 1 min. The flow through was discarded and the column was then washed with PE buffer[™] and residual buffer removed by further centrifugation. The plasmid was then eluted in 50 µl of molecular grade H₂O and stored at -20 °C.

2.1.8 Site directed mutagenesis of OBP3

Mutagenesis was performed using a QuikChange[®] Lightning Site-Directed Mutagenesis Kit (Agilent Technologies). The way this kit works is shown in figure 2.1.

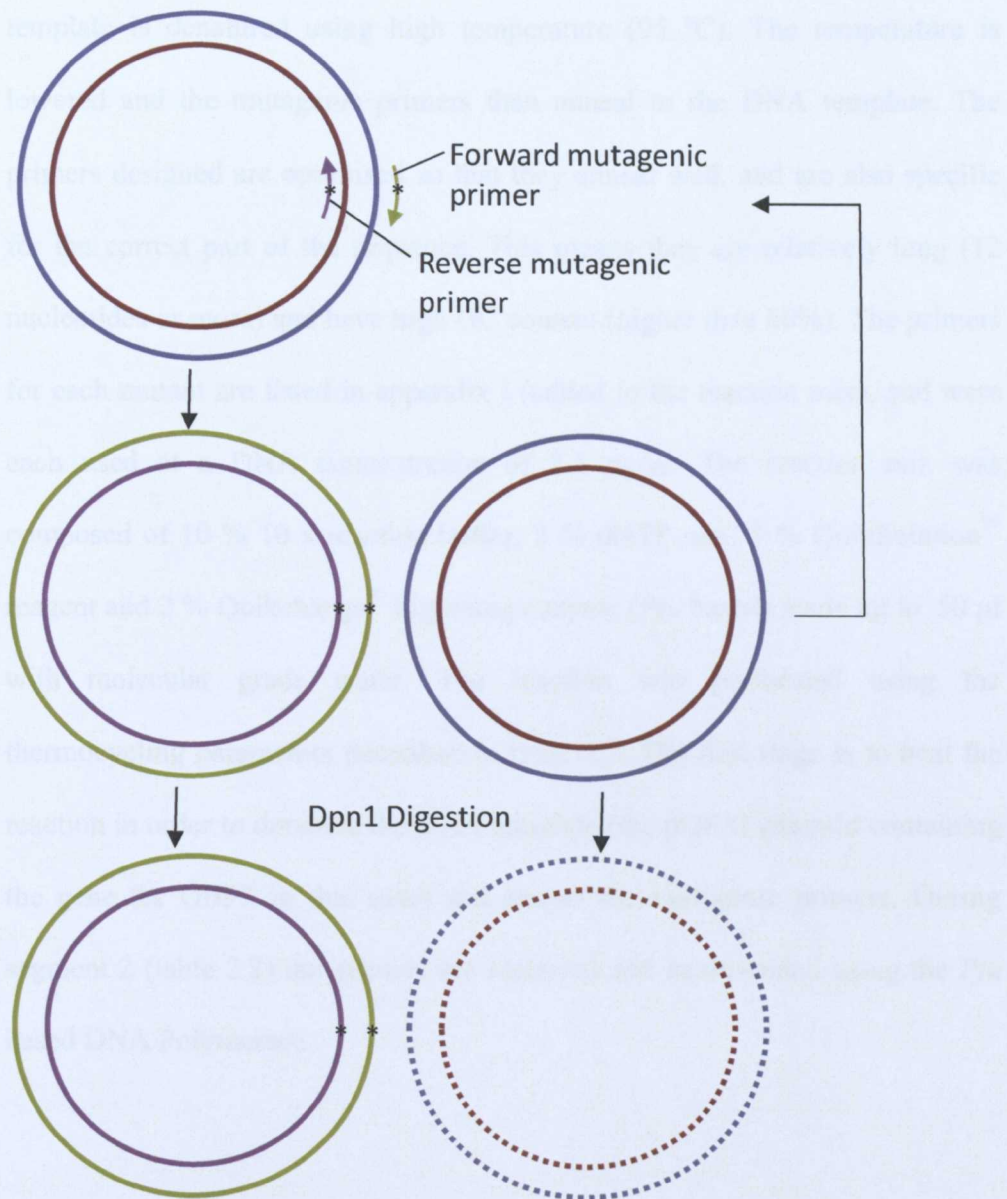


Figure 2.1 Mutagenesis using the QuikChange® Lightning Site-Directed Mutagenesis Kit. The DNA template is shown in blue (forward strand) and brown (reverse). The mutagenic primers are labelled and the mutation indicated by a * on each strand. The result of mutagenesis is shown in green (forward strand) and purple (reverse strand). Dpn 1 is used to digest the DNA template (dotted lines).

Firstly primers were designed and synthesised (Fisher, Loughborough, UK) to incorporate the desired mutation (indicated by * in figure 2.1). Firstly the DNA

template is denatured using high temperature (95 °C). The temperature is lowered and the mutagenic primers then anneal to the DNA template. The primers designed are optimised so that they anneal well, and are also specific for the correct part of the sequence. This means they are relatively long (12 nucleotides or more) and have high GC content (higher than 60%). The primers for each mutant are listed in appendix I (added to the reaction mix), and were each used at a DNA concentration of 2.5 ng/μl. The reaction mix was composed of 10 % 10 x reaction buffer, 2 % dNTP mix, 3 % QuikSolution™ reagent and 2 % Quikchange® Lightning enzyme (*Pfu* based) made up to 50 μl with molecular grade water. The reaction was performed using the thermocycling parameters described in table 2.2. The first stage is to heat the reaction in order to denature the DNA template (the pQE31 plasmid containing the gene for OBP3 in this case) and anneal the mutagenic primers. During segment 2 (table 2.2) the primers are extended and incorporated using the *Pfu* based DNA Polymerase.

The reaction was then treated with *DpnI* for 20 min at 37 °C. This enzyme digests methylated DNA (as produced by *E. coli*) and is therefore selective for just the original DNA template. Subsequently 2 μl of *DpnI* treated product was added to 45 μl ultracompetent XL10-gold cells which had been pre-treated with 5 % β-mercaptoethanol. A heat shock transformation was then performed in the same manner as described for M15 cells. Transformation into competent *E. coli* allowed the nicked plasmid to be repaired. The cells were spread on LB agar plates containing ampicillin (100 μg/ml) to select for the pQE31 plasmid.

The plasmid was later isolated as described earlier and sent for sequencing to check the correct mutation was present.

Segment	Cycles	Temperature	Time
1	1	95 °C	2 minutes
2	18	95 °C	20 seconds
		60 °C	10 seconds
		68 °C	2 minutes
3	1	68 °C	5 minutes

Table 2.2 Thermocycling parameters used for mutagenesis reactions

2.1.9 Sequencing

All plasmid DNA sequencing was carried out by Source Bioscience, Nottingham. Purified plasmid was eluted in molecular grade H₂O. The concentration was checked using a Nanodrop™ ND1000 spectrometer (Thermo Fisher Scientific, Wilmington, USA) by measuring the absorbance at a wavelength of 260 nm.

2.2 Production of Rat OBP3

2.2.1 *E. coli* expression systems used

The genetics and molecular biology of *E. coli* are well understood through many years of study and therefore the genetic manipulation required to produce heterologous protein is relatively simple. Proteins not originally derived from *E. coli* are usually well tolerated. Additionally, the growth medium required is low-cost and cultures are easily grown. Protein yields are reasonable, though not as high as those reported for yeast systems (Sambrook & Russell 2001) which are also capable of post-translational modifications (such as glycosylation) which *E. coli* is not. However, *E. coli* is more suited for isotopic labelling, which is required for Nuclear Magnetic Resonance studies (Chapters 4 - 6) as the supplements and media used are more straightforward and inexpensive.

2.2.2 Plasmids

The pQE31 plasmid (Qiagen) containing the *obp3* gene, and chemically competent *E. coli* M15 were kindly donated by Loic Briand (INRA, France). The *obp3* gene had previously been cloned into the pQE31 vector (Qiagen) by the Breer laboratory (Hohenheim, Germany). The pQE plasmids were originally derived from the pDS56/RBSII and pDS781/RBSII-DHRS plasmids (Stüber et al., 1990). RBSII is a synthetic ribosomal binding site (RBS) included to enable high translation rates. Additionally pQE plasmids contain the coliphage T5 promoter, which is recognized by the *E. coli* RNA

polymerase. This requires the *lac* repressor, constitutively expressed by the *lacI* gene on the pREP4 plasmid, to achieve tight regulation. (Farnbaugh, 1978). Figures 2.2 and 2.3 show pQE31 and pREP4 respectively, highlighting their key features.

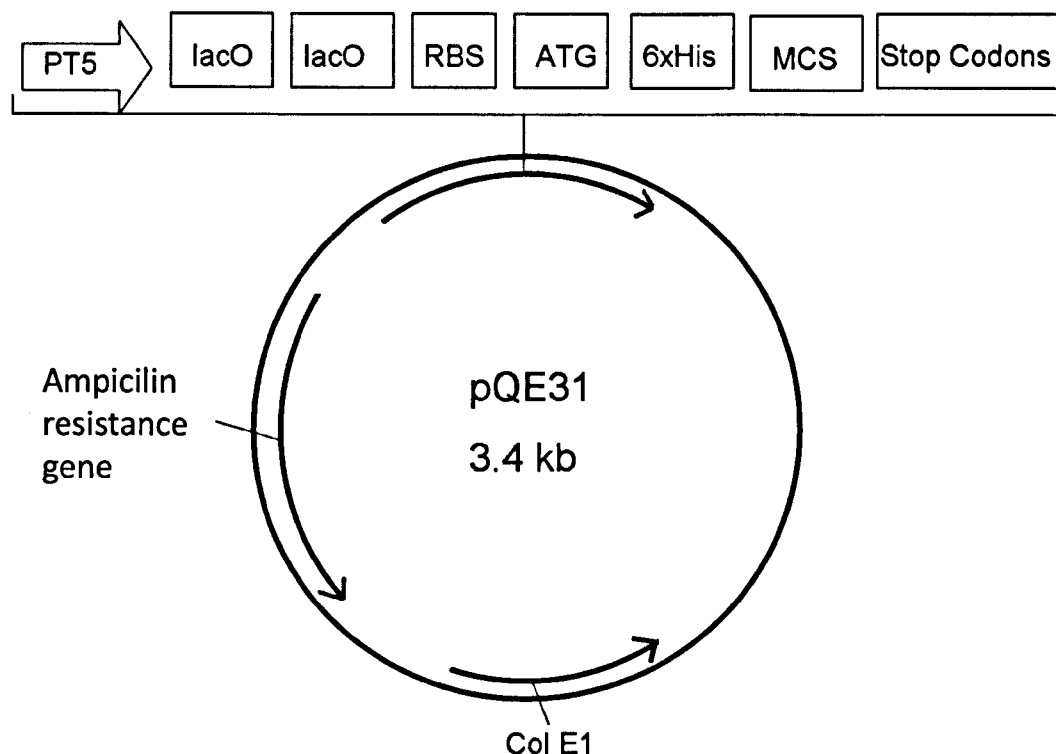


Figure 2.2 The pQE31 plasmid (Qiagen). The gene encoding OBP3 is inserted between the NdeI and BamHI restriction site of the Multiple Cloning Site (MCS). The protein is histidine tagged at the N-terminus. PT5 – T5 promoter, lacO - lac operator sequence, RBS – ribosome binding site, ATG – start codon, 6xHis – histidine tag. The plasmid carries the ampicillin resistance gene to allow selection of the plasmid.

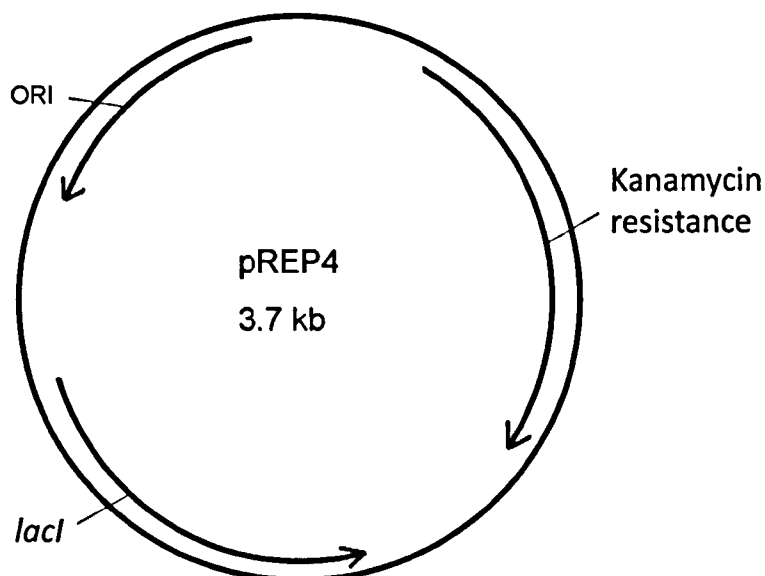


Figure 2.3 The pREP4 plasmid contains the constitutively expressed *lacI* gene which represses the T5 promoter on pQE31. Induction by isopropyl β -D-1-thiogalactopyranoside relieves this repression. The plasmid is selected by the presence of kanamycin resistance.

2.2.3 OBP3 Sequence

A six-histidine tag was included at the N-terminus of OBP3. An additional four residues before and three residues after the histidine residues were included, as shown in figure 2.4 along with the original sequence from *Rattus norvegicus* with the export signal sequence highlighted in orange.

Amino acid sequence of full length OBP3 (including signal sequence)

MRGSHHHHHHTDAEEASFERGNLDVDKLNQDWFSIVVASDKREKIEENGSMRVFVQH
IDVLENSLGFTFRIKENGVCTEFSLVADKTAKDGEYFVEYDGENTFTILKTDYDNYV
MFHLVNVNNGETFQLMELYGR TKDLSSDIKEKFAKLCVAHGITRDNIIDLTKTDRCL
QA

Amino acid sequence of the histidine-tagged OBP3 construct

MRGSHHHHHHTDPEEASFERGNLDVDKLNQDWFSIVVASDKREKIEENGSMRVFVQH
IDVLENSLGFTFRIKENGVCTEFSLVADKTAKDGEYFVEYDGENTFTILKTDYDNYV
MFHLVNVNNGETFQLMELYGR TKDLSSDIKEKFAKLCVAHGITRDNIIDLTKTDRCL
QA

Figure 2.4 Amino acid sequence of OBP3 prior to export into the extracellular space from the cytoplasm, where the export signal, shown in orange, is cleaved. The recombinant OBP3 construct used in this study has this sequence replaced by a six histidine tag (shown in red) and an additional seven residues (shown in blue). Throughout this thesis OBP3 is numbered such that residue 1 is the first residue after the engineered tag shown in blue and red. For the purpose of NMR assignment the tag is numbered using negative values, with the proline at the end of the tag being residue 0.

The inclusion of the histidine tag allowed OBP3 to be subjected to metal affinity chromatography as part of a purification scheme, to produce sufficient protein which was suitably pure for biophysical studies.

2.2.4 Overexpression of non-isotopically labelled OBP3 in *E. coli*

E. coli M15 cells containing both pREP4 and pQE31 were used to overexpress OBP3 protein. Standard growth conditions, including ampicillin and kanamycin, were used to produce four 25 ml overnight cultures from single colonies of *E.coli* M15/pREP4/pQE31 and these were used to inoculate four 250 ml flasks of fresh LB medium. The new cultures were grown to an

OD_{600nm} of 0.6 before protein production was induced by adding Isopropyl β -D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. After a further 4 hours of growth the cells were harvested at 4 000 g for 20 min.

2.2.5 Overexpression of isotopically labelled OBP3 in *E. coli*

To overexpress isotopically labelled OBP3 for NMR studies in *E.coli*, the same procedure was used, except that the overnight cultures were centrifuged at 2500 g to pellet the cells which were used to inoculate a total of 1 l of defined medium (6 g/l Na₂PO₄, 3 g/l KH₂PO₄, 0.5 g/l NaCl, 0.3 g/l MgSO₄, 15 mg/l CaCl₂, 1 g/l NH₄Cl, 4 g/l C₆H₁₂O₆, 10 mg/l biotin and 10 mg/l thiamine) containing ampicillin and kanamycin. For ¹⁵N labelling the ammonium chloride was replaced with ¹⁵N ammonium chloride (Cambridge Isotope Laboratories, Massachusetts, USA) and for ¹³C labelling the glucose was replaced with 2 g/l ¹³C₆ glucose (Cambridge Isotope Laboratories). Again the cells were grown to an OD_{600nm} of 0.6 before protein production was induced by adding Isopropyl β -D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. At this point the temperature was dropped to 25 °C and ampicillin (100 μ g/ml) added to replace that which had been used up by the *E.coli*. Growth was continued for 16 hours before the cells were harvested at 4000 g for 20 min. Extraction and purification were carried out in the same manner for protein with or without an isotopic label.

2.2.6 Selective non-isotopic labelling of OBP3

Here the same procedure used for isotopically labelling OBP3 was used to selectively “unlabel” (or more accurately, not label) specific residues, such as lysine. The only alterations were that 0.5 g/l of the desired ^{14}N amino acid was added to the initial defined medium followed by a further 0.5 g/l after induction.

2.2.7 Extraction of OBP3

The harvested pellet after overexpression was resuspended in 20 ml of lysis buffer (50mM sodium phosphate, 300 mM NaCl, 20 mM imidazole, 1 mM phenylmethylsulphonyl fluoride (PMSF), pH 8.0) and incubated on ice for 10 min. The cells were then sonicated on ice using a 9.5 mm solid probe fitted to a Soniprep 150 (MSE, London, UK) with an output frequency of 23 KHz. 15 cycles of 20 s sonication at an amplitude of 10 microns followed by 40 s of recovery were performed. The lysate was then centrifuged at 30 000 g for 30 min at 4 °C. The supernatant was decanted to a fresh tube and the centrifugation repeated to ensure all cell debris had been removed. The supernatant was then passed through a 0.2 µm syringe filter (Sartorius, France) recovered and stored at 4 °C ready for purification.

2.3 Purification of OBP3

2.3.1 Metal Affinity Chromatography

A HisTrap™ HP Column packed with Ni Sepharose (1.6 i.d., 5 ml volume, GE Healthcare) was used for purification as OBP3 produced from the pQE31 plasmid was His-tagged at the N-terminus (Qiagen Expressionist manual). The column was attached to an ÄKTA™ Purifier system. The column was washed with 5 column volumes of binding buffer (50 mM sodium phosphate, 300 mM NaCl, 20 mM imidazole) prior to the filtered lysate being applied to the column by hand at an approximate flow rate of 1 ml.min⁻¹. The column was washed with 15 column volumes of binding buffer to remove unbound protein and cell debris, after which the protein was eluted by increasing the imidazole concentration to 300 mM in a single step, at which point the flow rate was reduced to 0.5 ml.min⁻¹. 1 ml fractions of flow through were collected. The presence and purity of OBP3 was confirmed by SDS-PAGE.

2.3.2 Gel Filtration

A Superdex 75 gel filtration column (24 ml bed column, 1.0 i.d. x 75 cm) was conditioned with 30 mM potassium phosphate, 100 mM NaCl, pH 7.0. Subsequently 5 ml of sample was injected onto the column via a 10 ml sample loop. If necessary the protein was first concentrated to reduce the volume to below 5 ml using a VivaSpin centrifugal spin concentrator (GE Healthcare). Elution was carried out at 1 ml.min⁻¹ using the same buffer. The resultant

fractions were pooled prior to endogenous ligand removal and small aliquots run on an SDS-PAGE gel.

2.3.3 Endogenous Ligand Removal

In order to remove ligands that may have already been present in the OBP3 binding pocket, the protein was ethanol precipitated by the method of Bingham et al. (2004). The resultant fractions from gel filtration were pooled before being concentrated to a volume approximately 10 times smaller. To a protein solution of volume ' χ ', 2 χ volumes of 100% analytical grade ethanol (Fisher, Loughborough, UK) was added. The solution was incubated at 4°C with gentle shaking for 2 hours after which time the solution was centrifuged at 5 000 g for 15 minutes to pellet the protein. After the supernatant was discarded the pellet was lyophilised with a bench top -80 °C freeze dryer (Telstar cryodos-80, Progen, London, UK) in order to remove any remaining ethanol.

2.3.4 Desalting

The freeze dried pellet was resuspended as fully as possible in MilliQ® (Millipore, Massachusetts, USA) H₂O, any insoluble material was removed by centrifugation at 4000 g for 10 min. The protein solution was then applied to a HiTrap™ desalting column (25 ml G25 column, GE Healthcare), pre-equilibrated with MilliQ® H₂O. For each run 2.5 ml aliquots were injected and subsequently eluted with the MilliQ® H₂O. Protein free of salt was then frozen

at -80 °C, prior to being lyophilised. The lyophilised protein samples were stored at -20 °C.

2.4 Assessment of protein and purity and quantity

2.4.1 Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Protein purity and approximate molecular weight was assessed using 12 % denaturing gels, the composition of which is shown in table 2.3. All SDS-PAGE was performed using a BioRad Protean® III system. The manufacturer's instructions (BioRad, Hercules, CA, USA) were followed for casting gels and setting up the tank. The inner chamber was filled with cathode buffer (250 mM Tris-HCl), 250 mM Tricine, 0.25 % SDS, pH 8.25) and the outer chamber with anode buffer (400 mM Tris-HCl, pH 8.8 (no SDS)).

Component	Stacking Gel (µl)	Resolving Gel (µl)
Bisacrylamide:Acrylamide (19:1)	250	950
Gel Buffer (3 M Tris, 0.3 % (w/v) SDS, pH 8.45	620	1250
Distilled Water	1610	1120
Glycerol	0	400
N,N,N',N' Tetramethylethylenediamine	2	2
10 % Ammonium Persulphate	20	19

Table 2.3 Composition of SDS-PAGE 12% gels

Samples for SDS-PAGE were prepared by mixing protein solution 1:2 with reducing buffer (BioRad Tricine Sample Buffer) in order to reduce disulphide bonds in the protein. This was then heated for 5 min at 95 °C to denature the protein, briefly centrifuged, and then loaded onto the gel. A standard protein marker (#7720, Broad Range pre-stained marker, New England Biolabs) was used to enable assessment of the size of the protein.

An initial running voltage of 60 V was used which was increased up 150 V once the stacking layer had been passed. Once the gels had run they were removed and stained for one hour using Coomassie Blue powder (1% w/v) in 50 % (v/v) water, 40% methanol (v/v), 10 % acetic acid (v/v) and then destained for one hour in a solution of the same composition but without Coomassie Blue.

2.4.2 Mass spectrometry

For native mass spectrometry lyophilized protein was dissolved in 25 mM ammonium acetate to a concentration of 1 mg/ml and injected into an Electrospray Ionisation Mass Spectrometry at infusion rate of 5 $\mu\text{l min}^{-1}$ (SYNAPT™ Waters, Hertfordshire, UK). For denaturing mass spectrometry, the same procedure was followed except lyophilized protein was dissolved in 50:50 mixture of methanol and ammonium acetate (to give a final concentration of 25 mM) spiked with 1% acetic acid. All data were analysed using MassLynx™ software (Waters).

2.4.3 Determination of protein extinction coefficients

In order to determine the quantity of protein in solution for subsequent biophysical studies the extinction coefficients were determined by the method of Gill and von Hippel (Gill and von Hippel, 1989). Freeze dried protein was dissolved in denaturing solution (6 M guanidine hydrochloride, 100 mM dibasic sodium phosphate, pH 6.5). This was used to make 1 in 10 dilutions in both denaturing solution and a renaturing solution (100 mM dibasic sodium phosphate, pH 6.5). The absorbance of these dilutions was determined at a wavelength of 280 nm. The theoretical extinction coefficient (ϵ_{280}) was calculated from the protein sequence (Gill and von Hippel, 1989) and then the value ($12330 \text{ M}^{-1}.\text{cm}^{-1}$) used in the following calculations shown in equation 2.1 to give a corrected extinction coefficient of $13454 \text{ M}^{-1}.\text{cm}^{-1}$. This was an increase of 9.1 %. To calculate the concentration of mutants of OBP3, the theoretical extinction coefficient of the mutated sequences was increased by 9.1% to give a more accurate extinction coefficient.

$$\frac{A_{280\text{nm unfolded}}}{\text{Theoretical } \epsilon_{280}} = [\text{unfolded}]$$

$$\frac{A_{280\text{nm 1:10 dilution of refolded protein}}}{0.1[\text{unfolded}]} = \text{Corrected } \epsilon_{280} (\text{M}^{-1}\text{cm}^{-1})$$

(Equation 2.1)

2.5 Biophysical and structural techniques

2.5.1 Isothermal titration calorimetry experiments

The OBP – odour binding system is somewhat unusual as odours are volatile and thus enter the nasal cavity in the gas phase. They then enter the nasal mucus, which is a hydrophobic, viscous liquid and are conferred to the olfactory receptor neurons (as described in section 1.1.3). Attempts have been made in our laboratory to study this transition, and the role of OBPs within the nasal mucus by constructing near physiological replica systems (Taylor et al. 2008; Yabuki et al. 2010, 2011; Borysik et al. 2010). ITC is an entirely solution based technique and therefore does not reflect the complete *in vivo* situation but mimics the aqueous odour-OBP interaction. It is a more valuable technique than physiological-like techniques for giving a full thermodynamic characterisation of the interactions with direct calculation of binding affinities.

The odours used were liquid at room temperature, this made calculation of the exact concentration a little more complicated. The exact method used is described in section 2.6.2.

It is standard practise to degas all samples and buffers thoroughly before an ITC experiment to reduce heat changes involved in the mixing of air contained in the solutions. Due to the volatility of the odour molecules this posed a significant problem as degassing of odour solutions led to an alteration of the number of odour molecules in solution and hence an uncertainty as to the exact

ligand concentration. Therefore the buffer was thoroughly degassed before the odour was added. The effect of dissolution of the odour solution into the protein solution was accounted for as fully as possible using a control experiment titrating odour solution into the protein buffer alone, the data from which was subtracted from the data from the experiment of odour titrated into OBP3 experiment.

Ideally the use of plastic should have been avoided throughout the ITC experiment as plasticisers very much resemble odour molecules and were likely to readily bind to the protein. This was not possible for the 96 well loading tray for the ITC instrument but potential problems were minimised by thoroughly washing the tray before use. Additionally only one sample was loaded at a time, therefore contact time with the plastic and the amount of binding to the tray material was both minimised and consistent for all experiments.

All experiments were initially run using 30 μM protein solution, titrated with 300 μM ligand as this produced a heat change that was large enough to measure and binding curve transition from which a binding affinity could be determined. Concentrations were adjusted according to the results for ligands that bound more weakly. Ligands and protein were diluted in 10 mM Phosphate Buffered Saline (PBS), pH 7.4 (8 g/l NaCl, 0.2 g/l KCl, 0.24 g/l KH_2PO_4 , and 1.44 g/l Na_2HPO_4). ITC experiments were performed using a Microcal VP-Auto ITC (GE Healthcare). Every 180 s 14.3 μl of ligand solution

was injected into the protein solution in the calorimetric cell. Throughout the titration the solution was stirred at 300 rpm. A preliminary injection of 2 μ l was carried out; this data point was routinely deleted from each spectrum. As the injection was small relative to the subsequent injections the effect on curve fitting was negligible. This preliminary injection is necessary to negate the effect of slow diffusive exchange from the ligand solution in the syringe tip to the calorimetric cell which may take place during the relatively long pre-equilibration period prior to injections.

2.5.2 Preparation of odorants for ITC experiments

The initial concentration of a neat odorant (or ligand), IC was determined by the formula in equation 2.2.

$$\frac{LD}{MW} \times P = IC$$

(Equation 2.2)

Where LD = ligand density in g/l, MW = molecular weight and P = ligand purity.

Ligands were diluted in the required aqueous buffer to make a stock (preferably at ~ 5 mM). Some of the odorants used were only slightly soluble in aqueous buffer. Organic solvents, particularly methanol should be avoided where possible in ITC experiments as large heat changes take place when they are dissolved in water. Previous studies have used methanol as an initial solvent, followed by adjustment of the final methanol concentration to below

1% (Löbel et al. 2002; Borysik et al. 2010). Where possible dilute solutions were made up which did not require methanol to solubilise the ligand. When this was not possible it is indicated in appendix II. When preparing ligands only glass apparatus was used, and odorants were stored in bottles with air tight metal lids. This avoided the odour molecules being adsorbed onto the surface of the container and affecting the final concentration. Details of all ligands used and how they were made up can be found in the appendix II.

2.5.2.1 Fitting of Isothermal titration calorimetry data.

Titration curves were plotted using Origin 7.0 software (corrected by subtraction of the titration curve generated by injection of the same amount of ligand into a cell containing buffer alone). Experiments in this thesis were best fitted to a one site model, meaning that for each macromolecule one ligand molecule is bound. The model is described by the following equation:

$$\frac{\delta Q}{\delta [X_t]} = \frac{\Delta H \cdot V_0}{2} \left[1 + \frac{1 - \frac{[X_t]}{n[M_t]} - \frac{1}{n[M_t]K}}{\left\{ \left(\frac{[X_t]}{n[M_t]} \right)^2 - \frac{2 \cdot [X_t]}{n[M_t]} \left(1 - \frac{1}{n[M_t]K} \right) + \left(1 + \frac{1}{n[M_t]K} \right)^2 \right\}^{\frac{1}{2}}} \right]$$

(Equation 2.3)

Where $[X_t]$ is the total ligand concentration, $[M_t]$ is the total macromolecule, V_0 is the effective volume of the calorimetric cell and dQ is the stepwise heat change for each point during the titration. This is the standard one site model included in Origin 7.0 for ITC (Microcal, GE Healthcare). Fitting was done

iteratively; using a non-linear regression procedure with the model equation to find the best values for the parameters (n , K and ΔH°). (Freyer & Lewis 2008).

2.5.3 Nuclear Magnetic Resonance Spectroscopy

2.5.3.1 NMR Spectrometers

Nuclear Magnetic Resonance (NMR) experiments were carried out using a Bruker-600 MHz spectrometer fitted with a triple resonance probe operating at 600.13 MHz. Direct carbon experiments were carried out on a Bruker-500 MHz spectrometer fitted with a cryoprobe. All spectra were acquired using standard Bruker pulse sequences.

2.5.3.2 Sample preparation

For all experiments, except D₂O exchange experiments (where the protein was dissolved in 100% D₂O), lyophilised OBP3 was dissolved in NMR buffer (10 mM potassium phosphate pH 6.1, 10 % D₂O, 0.04 % sodium azide),

2.5.3.3 Acquisition and processing software

NMR data were acquired using the TopSpin software package (Bruker). The standard experimental parameters used, including; the number of points, spectral widths and number of scans used in the acquisition of the data in this study, are listed in appendix III. Processing was also done using the TopSpin software package. The standard processing protocol included making strip transformation of the data and zero filling in the indirect dimensions, cosine window functions were routinely used. The CCPNMR software package (Vranken et al., 2005) was used for analysis of processed NMR data and the production of figures.

2.5.3.4 Titrations of ^{15}N labelled OBP3 with 2-isobutylthiazole

^{15}N labelled OBP3 at a concentration of 1 mM in NMR buffer (section 2.6.3.2) at pH 6.1 was titrated with increasing amounts of 2-isobutylthiazole (0.1 mM steps were used), until OBP3 reached saturation. At each step a ^1H - ^{15}N TROSY spectrum was recorded using the parameters in appendix III.

The chemical shift perturbations (CSP) for each residue were calculated using the following equation:

$$CSP = \sqrt{(\Delta\delta^1H)^2 + \left(\frac{\Delta\delta^{15N}}{5}\right)^2} \quad (\text{Equation 2.4})$$

Where $\Delta\delta^1\text{H}$ and $\Delta\delta^{15}\text{N}$ are the changes in the proton and nitrogen-15 chemical shifts respectively.

2.5.3.5 Preparation of ligand saturated samples

For NMR experiments where ^{15}N labelled OBP3 was required saturated with ligand, NMR buffer (pH 6.1) was used to solubilise the ligand. 20 ml volumes of buffer were required in order to dissolve the γ -lactones at concentrations up to 1.2 mM, beyond which the ligands were no longer soluble. The same method was used for 2-isobutylthiazole although it was soluble at concentrations beyond 4 mM. A 2:1 ratio of ligand to protein was used to ensure saturation (2 mM ligand, 1 mM protein).

2.5.3.6 Deuterium exchange NMR experiments

For D_2O exchanged $^1\text{H} / ^1\text{H}$ NOESY and $^1\text{H} / ^1\text{H}$ TOCSY experiments lyophilized OBP3 protein was re-suspended in 100% D_2O and left at room temperature for 8 hours before being resubmitted to freeze drying. This process was repeated three times to ensure full exchange had taken place. Finally OBP3 in D_2O placed in a sample tube and standard $^1\text{H} / ^1\text{H}$ NOESY and $^1\text{H} / ^1\text{H}$ TOCSY experiments using the parameters in appendix III were conducted.

For the deuterium exchange $^1\text{H} / ^{15}\text{N}$ TROSY experiments (described in section 6.2.1.) lyophilised protein was dissolved in 100 % D_2O and immediately placed in the pre-shimmed 600 MHz spectrometer machine and a $^1\text{H} / ^{15}\text{N}$ TROSY with 2048 points in the direct dimension and 64 points in the indirect dimension, 8 scans per t_1 increment were conducted as described for the standard experiment described in appendix III.

2.5.3.7 $^1\text{H} / ^{15}\text{N}$ Heteronuclear-NOE NMR experiments

A 1 mM unbound sample of OBP3 was prepared as described in section 2.6.2.3. Samples saturated with γ -lactones were prepared as described in section 2.6.3.5. $^1\text{H} / ^{15}\text{N}$ heteronuclear-NOE (het-NOE) values were determined by recording a HSQC spectrum (for OBP3 a TROSY based version of the experiment was used (Zhu et al., 2000)) with ^1H saturation applied before each scan is acquired and recording the same spectra without ^1H saturation (appendix III) using a relaxation delay of 5 s and saturation delay of 1 s were used. The ratio of peak intensities was then calculated. The peaks in the OBP3 spectra were picked manually before the in-built function in CCPNMR Analysis (Vranken et al., 2005) was used to calculate the het-NOE values for each residue. Peak heights were used to calculate the peak intensities, rather than peak volume to avoid error that may have been introduced by the shape of some of the weaker peaks. Weak peaks were present because the het-NOE experiments have inherently low sensitivity. This is because the pulse sequence starts with equilibrium ^{15}N magnetization instead of equilibrium ^1H magnetization (which is about 10 fold stronger) (Renner et al., 2002).

2.5.4 Crystallisation trials of OBP3

Protein prepared for crystallisation trials was not submitted to lyophilisation. After gel filtration the resultant fractions containing the protein were pooled and then exchanged into the buffer required for crystallography using a Vivaspin concentrator (GE Healthcare) and centrifuged in a benchtop conical tube centrifuge (Heraeus® Biofuge® Stratos). The volume was reduced to as

low as possible (~2 ml) then the concentrator refilled with 20 ml of the required buffer and the concentration repeated. This cycle was repeated six times. Finally the absorbance at a wavelength of 280 nm was checked and the concentration adjusted to that required (usually 10 mg/ml). In order to screen conditions for crystallisation Corning CrystalEX sitting drop crystallisation trays (Hampton Research, California, USA) were filled with a variety of reservoir solutions. Commercial screens as well as laboratory designed screens were utilised, full details of which are shown in appendix IV. Protein samples were diluted 1:1 with reservoir solution (to give a final volume of 100 nl) and deposited on the sample wells of the trays. This was done using a mosquito[®] Crystal nanolitre pipetting robot (TTP Labtech, Melbourne, Hertfordshire, UK). The trays were then sealed with Crystal Clear Sealing Tape (Hampton Research), stored at 291 K.

2.5.5 Circular Dichroism

An Applied Photophysics (Leatherhead, UK) π^* 180 spectrometer was used for Circular Dichroism (CD) experiments. Into an optical quartz cuvette (Hella QS-110) 300 μ l of a 1 mg/ml protein solution (in 10 mM potassium phosphate pH 6.1) was loaded. Scans were recorded between 190 nm and 260 nm at 10 nm min⁻¹ with a 1 mm path length. A temperature of 298 K was used. A scan of buffer alone was subtracted from the protein curve. The raw data in millidegrees was converted to mean residue molar ellipticity ($\Delta\epsilon_{MR}$) using the formula:

$$\Delta \varepsilon MR = \frac{\Delta A}{C \times n \times l} \quad (\text{Equation 2.5})$$

Where ΔA is the differential absorbance of left (A_{LCP}) and right circularly polarised (A_{RCP}) light as measured by the spectrometer, C is the molar protein concentration, n is the number of residues in the protein and c is the molar concentration of the protein sample and the path length (l) was 0.1 cm. To determining melting temperatures using CD, the signal at 217 nm was monitored as the temperature was raised in 1 K intervals from 298 K to 368 K.

2.5.5.1 Fitting circular dichroism melting temperature data

Melting temperature data was fitted using the software program Kaleidagraph (Synergy, Pennsylvania, USA) using the following equations;

$$\Delta G = \Delta H(1 - T/T_m - T) + T \ln(T/T_m) \quad (\text{Equation 2.6})$$

$$K = \exp(-\Delta G/(RT)) \quad (\text{Equation 2.7})$$

$$\alpha = Keq/(1 + Keq) \quad (\text{Equation 2.8})$$

$$[\Theta]_t = \alpha([\Theta]_F - [\Theta]_U) + [\Theta]_U \quad (\text{Equation 2.9})$$

Where $[\Theta]_t$ is the observed ellipticity, $[\Theta]_F$ is the initial ellipticity (folded state) and $[\Theta]_U$ is the final ellipticity (unfolded state). Additionally, α is the fraction folded species, T is the temperature and T_m is the melting temperature.

2.5.6 Differential scanning calorimetry

Differential scanning calorimetry (DSC) was performed using a Microcal VP-DSC (GE Healthcare). The DSC was thermally conditioned by being continually heated and cooled from 278 K to 393 K for five cycles. Lyophilised OBP3 was resuspended in 10 mM KHPO₄ buffer (pH 7.0) at a concentration of 50 μ M, 1 ml of this solution was placed in the sample cell. Buffer alone was placed in the reference cell. The cells were then heated from 278 K to 393 K at a rate of 90 °C / hr and cooled again at a rate of 60 °C / hr. Melting curves were plotted using Origin 7.0 software (corrected by subtraction of a blank experiment with buffer in both the reference and sample cells).

2.6.6.1 Fitting differential scanning calorimetry data

Corrected and normalized melting curves were analysed using Origin 7.0 software; firstly the baseline was established and then the curve was fitted to a two state model using the Levenberg/Marquardt non-linear least-square method which is in-built in Origin 7.0.

2.6 Homology Modelling

Homology models of OBP3 were generated and assessed using Swiss-PdbViewer and the automated homology server Swiss-Model (Arnold et al., 2006; Kiefer et al., 2009). ClustalW2 was used for primary sequence alignment (Chenna, 2003).

3. OVEREXPRESSION, PURIFICATION AND CHARACTERISATION OF OBP3

3.1. Introduction

In order to understand the binding interactions of OBP3 a number of biophysical techniques were employed. Initially, Isothermal Titration Calorimetry (ITC) was used to determine which odour compounds bound to OBP3. In order to do this, large amounts of pure OBP3 were required; therefore recombinant OBP3 was produced using the gram-negative bacterium *Escherichia coli*. This chapter describes the overexpression and purification of OBP3 from *E. coli*, including optimisation of the previously published method of producing OBP3 (Löbel et al. 2001). The initial characterisation by mass spectrometry (MS), 1D Nuclear Magnetic Resonance (NMR), Circular Dichroism (CD) and Differential Scanning Calorimetry (DSC) of OBP3 is described. In addition, attempts were made to crystallise OBP3 for structural studies.

3.2 Results and Discussion

3.2.1 Overexpression of OBP3

OBP3 from rat was overexpressed using *E. coli* M15 cells (these cells are particularly tolerant of toxic proteins, but in this case, the OBP3 protein wasn't toxic). A build-up of protein was seen over the 4 hours after induction with IPTG. Figure 3.1 shows the overexpression band just below 25 kDa, corresponding to OBP3, which from its primary sequence, is calculated to have

a molecular weight of 19.9 kDa. The protein was expressed into the soluble fraction, so denaturing conditions were not required for purification

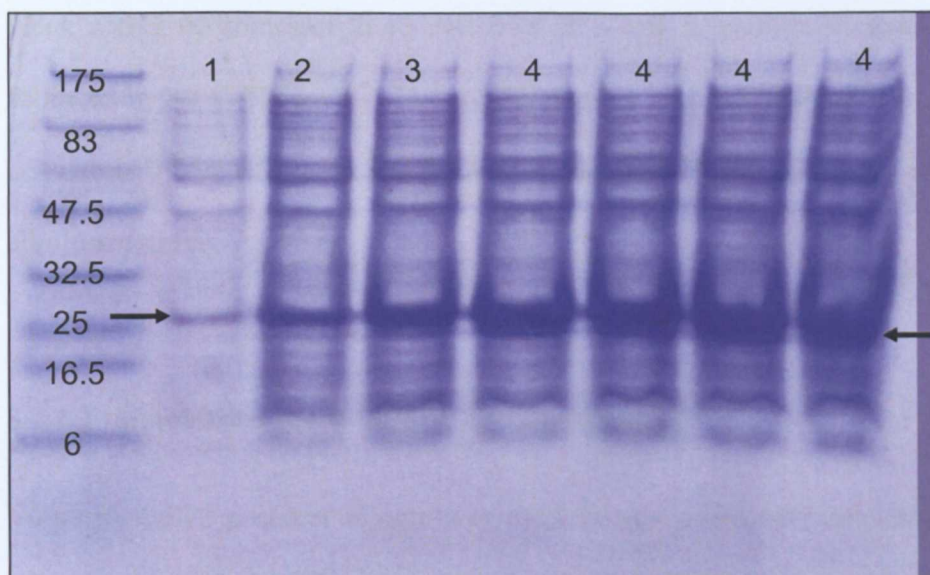


Figure 3.1. SDS-PAGE showing a accumulation of protein across four hours (lanes 1-4) of expression, including a sample of the final time point for all four flasks (all labelled 4 hours). The first lane shows the marker and the protein size the bands correspond to in kDa.

It was found that four hours growth after induction was sufficient to produce large amounts of protein from 1 litre of culture grown in LB broth.

3.2.2 Protein purification

3.2.2.1 Extraction

In order to purify OBP3 from *E. coli* M15 the cells needed to be lysed. This was done using sonication as per the method of Löbel et al. (1998, 2001); however some alterations to the original method were made. After harvesting the cells by centrifugation the pellet was stored at -20°C and thawed at room

temperature when required. Incubation with lysozyme on ice for one hour was omitted and instead phenylmethanesulphonyl fluoride (PMSF) was added immediately before sonication. The sonication was altered from a two minute block with a tip sonicator to 15 cycles of 20 s with a medium sonication tip, followed by 40 s of recovery. The soluble protein was then separated from the cell debris by centrifugation and the supernatant taken forward to metal affinity chromatography.

3.2.2.2 Immobilised metal affinity chromatography

To purify OBP3 a nickel column was used, however 300 mM imidazole was determined to be optimal to elute OBP3, rather than 250 mM used by Löbel et al. (2001). Figure 3.2 shows that a large number of *E. coli* proteins have been washed out of the column at the low imidazole concentration (20 mM) used during the wash step.

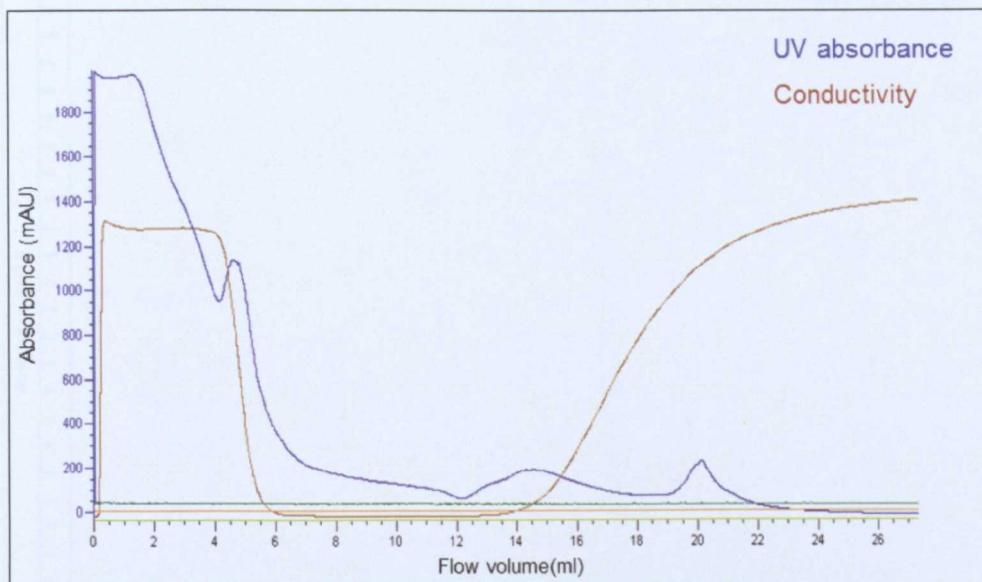


Figure 3.2 Chromatogram showing the UV absorbance (in blue) as the nickel column is washed with a low concentration of imidazole. Non-specifically interacting proteins are removed from the column after ~25 ml (5 column volumes), when UV absorbance returns to zero. Specific conductivity is shown in brown (measure in mS/cm) and changes to in response to the conductivity of the solution as it is eluted from the column and passes through the monitor. The remaining lines (orange and green) are unmonitored so remain at zero.

A one step elution method was used, whereby the imidazole concentration was increased to 300 mM immediately without the use of a linear or step-wise gradient. This resulted in a large single peak across six fractions of 1 ml in volume (between fractions 18-23), shown in figure 3.3.

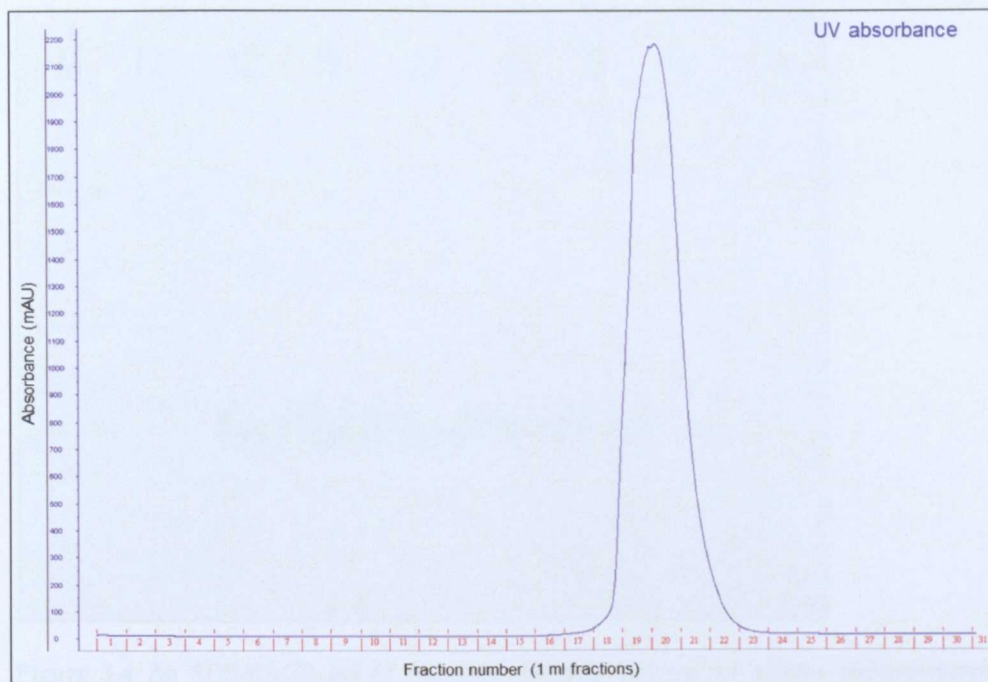


Figure 3.3 Chromatogram showing the elution of a large amount of protein between fractions 18-23 (6 ml), with a UV absorbance of almost 2000 mAU (an arbitrary value). The peak is somewhat asymmetric which may indicate the presence of contaminating proteins.

SDS-PAGE was used to assess the molecular weight, and purity. Figure 3.4 shows the SDS-PAGE gel and indicates a relatively clean band at just below the 25 kDa marker band, suggesting these fractions contain OBP3.

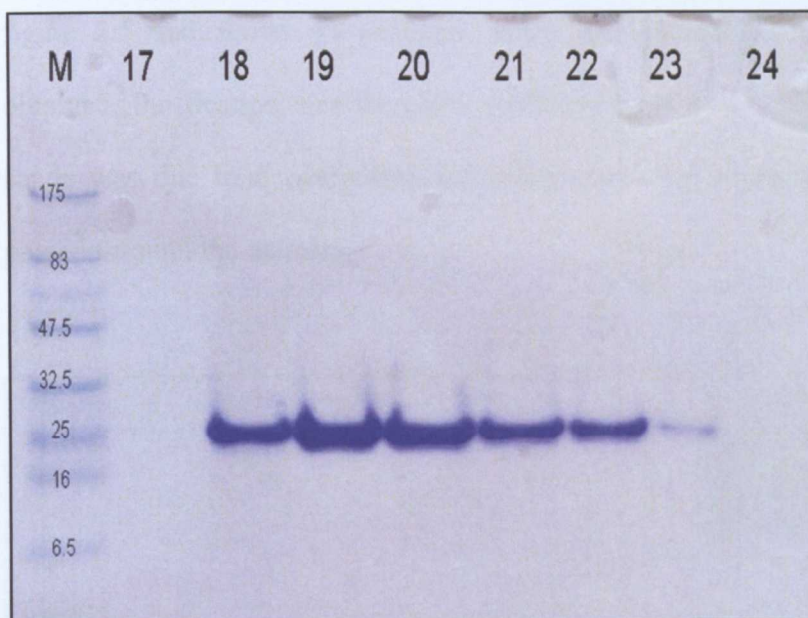


Figure 3.4 An SDS-PAGE gel of fractions resulting from nickel affinity chromatography showing apparently pure protein (at the concentration run on this gel) just below (25 kDa) in fractions 18-23. On the left is the molecular weight marker in kDa. The fraction numbers (corresponding to those shown in figure 3.3) are along the top.

3.2.2.3 Gel filtration chromatography

The fractions resulting from the affinity chromatography step were pooled together and injected onto the pre-equilibrated gel filtration column as described in section 2.3.2. Figure 3.5 shows that the largest peak is seen between fractions 14 and 19. There was a small contaminating peak between fractions 10 and 13. The major peak was somewhat asymmetric, appearing to tail off over a number of fractions later in the peak. SDS-PAGE of these fractions did not show a visible contaminating protein or fragmentation of the protein. To discover if the asymmetry was instead caused by a conformational effect or a protein-column interaction fraction 15, a relatively early fraction, was resubmitted to the column. The resulting trace is shown as an inset in

figure 3.5 and shows an identical, albeit smaller, peak to that originally obtained. Purification was therefore continued as it was concluded the peak shape was due to a conformational effect occurring when the protein was passed through the column.

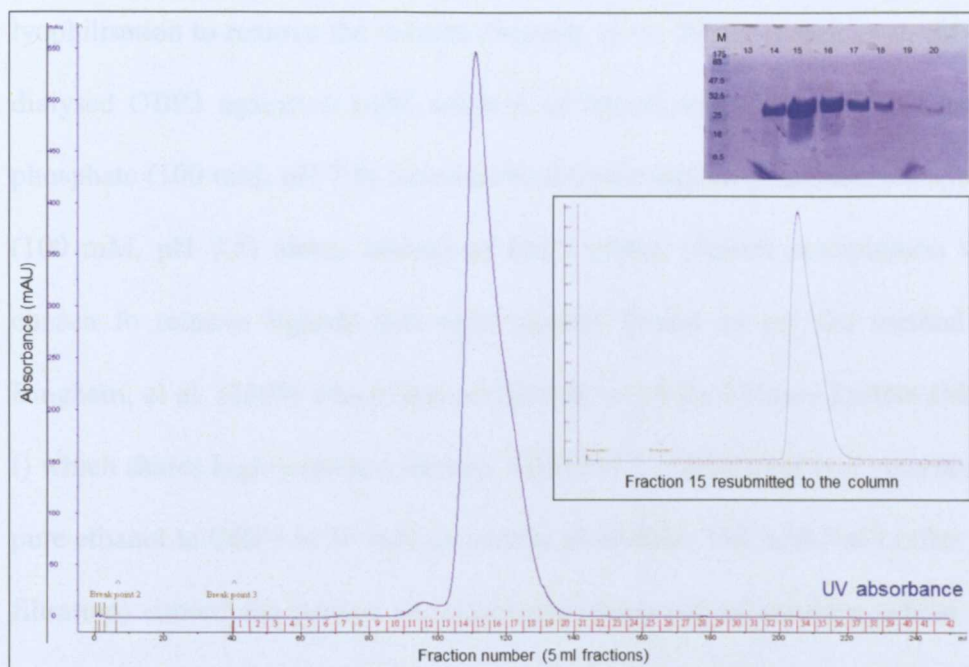


Figure 3.5 Chromatogram showing the elution profile of gel filtration chromatography (Superdex 75) of fractions 18-23 from the nickel affinity column (shown in figure 3.3). The UV absorbance (blue) shows an asymmetric peak. The top inset panel shows an SDS-PAGE gel of the fractions, indicating no obvious contamination (on the left is the marker in kDa with the fractions along the top). The bottom inset panel shows fraction 15 from the main peak resubmitted to the column. The peak shapes are very similar, suggesting the asymmetry may be due to a column effect rather than a contaminating protein.

3.2.2.4 Ligand removal

Odour binding proteins are promiscuous binders and readily bind ligands into their binding pockets (Tegoni et al. 2000). Through the stages of expression

and purification there is ample opportunity for this to occur. Indeed work on other odour binding proteins has reported the presence of ligands in the protein binding pocket after isolation (Bianchet et al., 1996; White et al., 2009). Different methods have been employed to try to achieve the apo-protein, for example dialysis in acetonitrile to unfold the protein, followed by lyophilisation to remove the solvent (Borysik et al., 2010). Löbel, et al. (2002) dialysed OBP3 against a 1mM solution of the odorant octanal in potassium phosphate (100 mM, pH 7.5) followed by dialysis against potassium phosphate (100 mM, pH 7.5) alone. Instead of this method, ethanol precipitation was chosen to remove ligands that were already bound as per the method of Bingham, et al. (2004) which was performed on Major Urinary Protein (MUP I) which shares high sequence identity with OBP3. Addition of two volumes of pure ethanol to OBP3 in 30 mM potassium phosphate, 100 mM NaCl (after gel filtration) caused the protein to unfold and come out of solution, whilst the ligand should remained soluble in the ethanol. Centrifugation was used to separate the protein away from the ligand-containing supernatant and lyophilisation was used to remove any traces of ethanol. The protein pellet was then resuspended in the desired buffer (usually water ready for desalting) (detailed method in section 2.3.3).

3.2.2.5 Desalting chromatography

Desalting effectively separates salt from the protein which is essential for NMR. After ethanol precipitation it appeared there was still some salt remaining in the protein pellet. The pellet was resuspended in water and the G25 desalt column was also conditioned with water. Figure 3.6 shows the

separation of OBP3 from the remaining salt. A small peak before the larger peak is likely to be a small amount of aggregation of the protein on the column, possibly due to the effect of being in pure water.

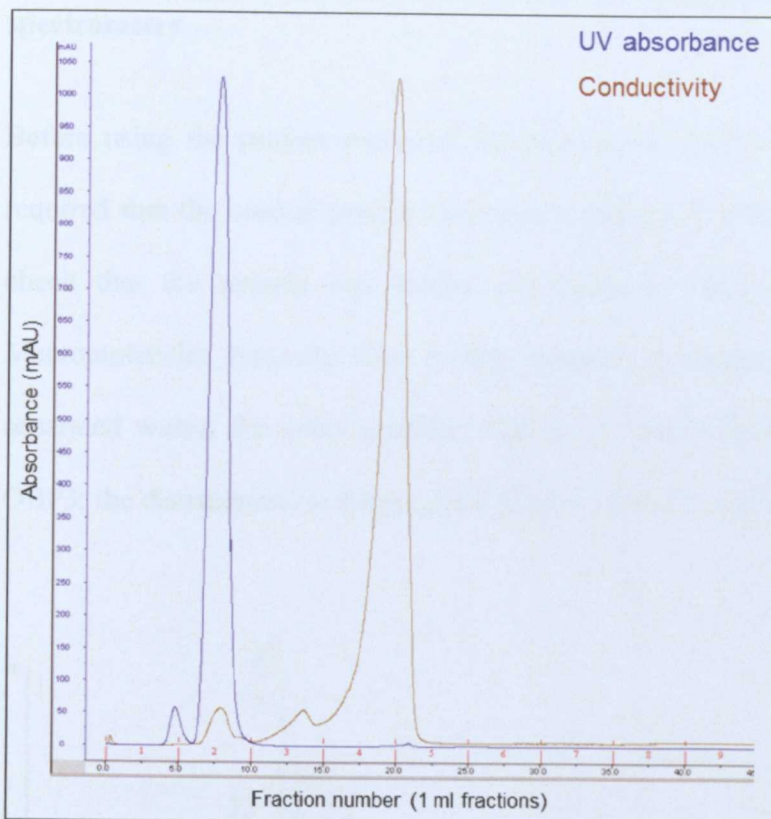


Figure 3.6 Desalting chromatography using a G25 column. In blue is the UV signal from the eluted protein, whilst in brown is the conductivity reading, indicative of the salt in the sample. Clear separation is seen. The small peak to the left of the main peak may indicate a small amount of aggregation due to the protein being in pure water.

Prior to use, protein that had been desalted, was lyophilised and stored at -20°C. This was resuspended in the desired buffer immediately before being used for biophysical studies, ensuring a low chance of ligands binding to the OBP3 before experiments were carried out. Average yields of 50 mg per litre

of bacterial culture have been achieved, which is an improvement of 32 mg above the yield quoted by Löbel, et al. (2002).

3.2.3 Determination of the molecular weight of OBP3 by mass spectrometry

Before using the protein produced for biophysical studies confirmation was required that the correct protein had been produced. It was also necessary to check that the protein was folded and therefore likely to be functional. Macromolecules typically have a large number of charge states. These are separated within the mass analyser. Figure 3.7 shows the mass spectrum of OBP3; the distribution of charge states from 9 + to 18 + can be seen.

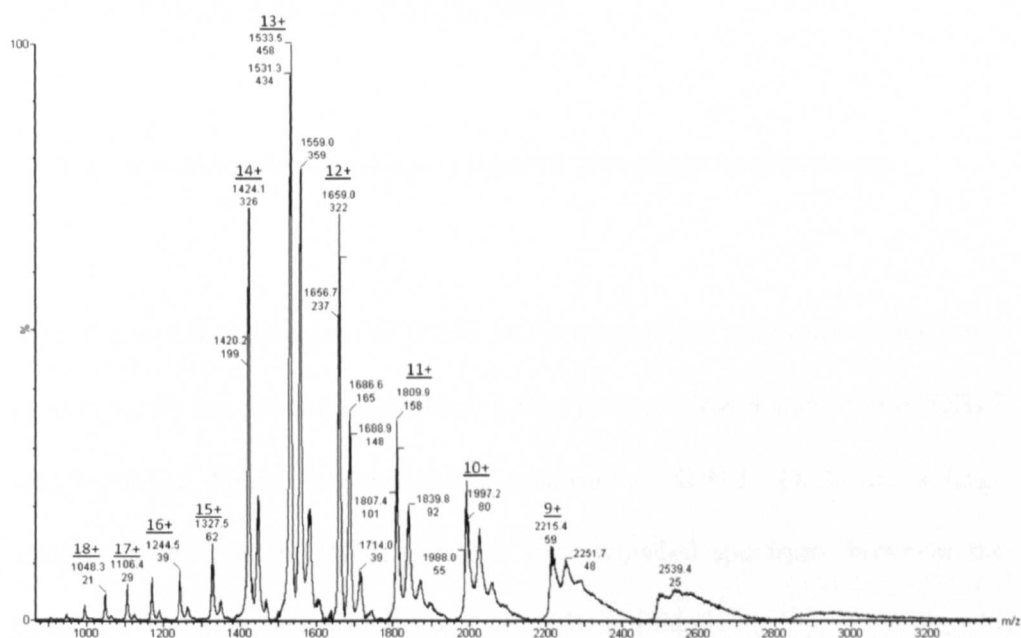


Figure 3.7 The mass spectrum of OBP3 in denaturing conditions. Analysis of the charge states gives a molecular weight of 19 897 kDa for the largest peak, which is equal to the molecular weight calculated from sequence.

It was noted that for each charge state there appeared to be a major peak with a molecular weight of 19 897 kDa, which corresponds exactly to the calculated molecular weight of OBP3 using monoisotopic masses. There were also minor peaks giving molecular weights of 19 869 kDa (28 Da less than the major peak) and 19 926 kDa (29 Da more than the major peak). It is unclear why these peaks exist, as they do not correspond to an obvious mutant of OBP3. It may be the case that the lower number corresponds to OBP3 losing its N-terminal methionine, but with the addition of an artefact of mass spectrometry (such as phosphate ions), whilst the higher molecular weight represents OBP3 plus a mass spectrometry adduct. Analytical ultracentrifugation was therefore used to confirm a single species was present in solution (Rhodes et al., 2009).

3.2.4 Assessment of the structure of OBP3

3.2.4.1 One-dimensional nuclear magnetic resonance spectroscopy

After the initial purification of OBP3, nuclear magnetic resonance (NMR) was used to verify the protein was folded. Using proton NMR a spectrum of OBP3 was recorded. Figure 3.8 shows the spectrum of OBP3. There are a large number of protons in OBP3, giving a very crowded spectrum, however the peaks are well dispersed which suggests OBP3 is folded.

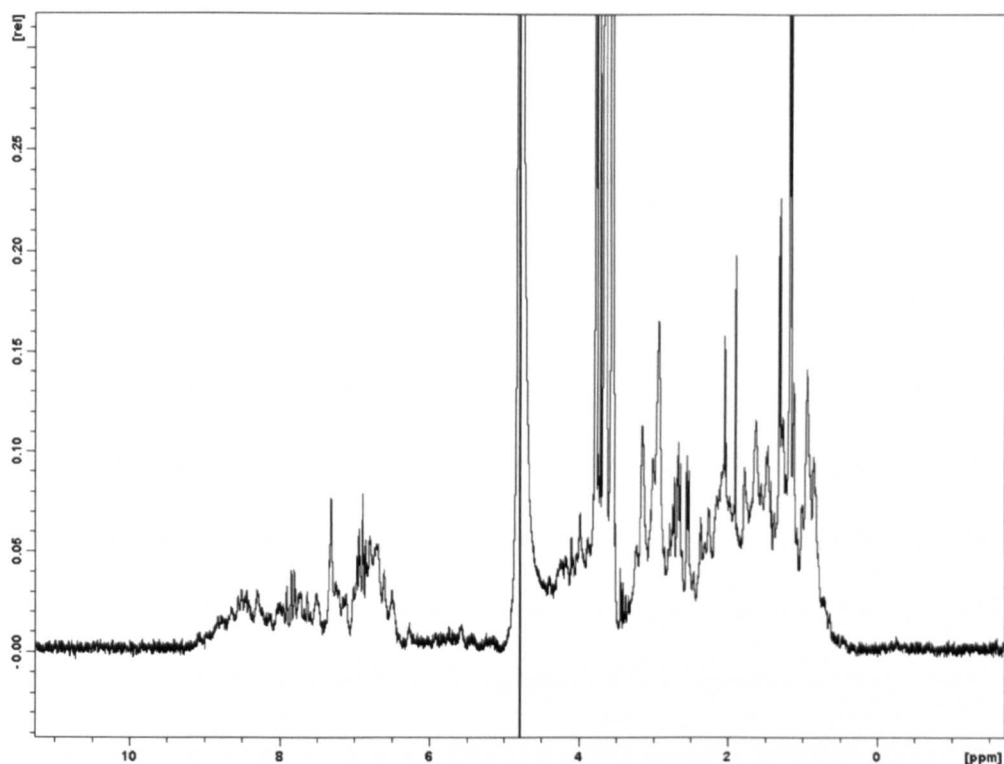


Figure 3.8 1D NMR spectrum of OBP3 recorded in 10 mM Potassium phosphate, 10 % D₂O, 0.04 % sodium azide, pH 7.0. A large dispersion of chemical shifts suggests that the protein is folded. It also appears that there aren't any small molecular weight contaminants.

3.2.4.2 Using circular dichroism to determine the secondary structure content of OBP3

Circular dichroism (CD) is a solution based technique, which can be applied to estimate the secondary structure composition of a protein sample.

The CD spectrum of OBP3 is shown in figure 3.9. Measurements were not made below 190 nm due to the limitations of the instrument used. Ideally data points should be recorded down to 170 nm and below, as can be recorded using synchrotron radiation, in order to reliably estimate the secondary structure of the protein (Corrêa & Ramos, 2009). From the spectrum the abundance of anti-

parallel β -sheet is clearly seen by the presence of a negative curve between 205 – 240 nm. The amount of α -helix is more difficult to ascertain due to the lack of data in the lower part of the spectrum, but the relatively sharp positive curve between 190 nm and 205 nm indicates some helical presence.

When the data were submitted to Dichroweb in order to determine a quantitative measure of the secondary structure, somewhat unexpectedly the protein appears to have over 63 % helical content, 20 % random coil and only 17 % β -sheet. These structural predictions do not concur with the expected secondary structure of OBP3, as a member of the lipocalin protein family, having eight anti-parallel β -sheets and an α -helix at the C terminal end of the protein. This may be due to a lack of good quality data at the lower wavelengths of the spectra which has distorted the fitting and thus the secondary structure predictions. NMR will be used to predict the secondary structure in addition to CD.

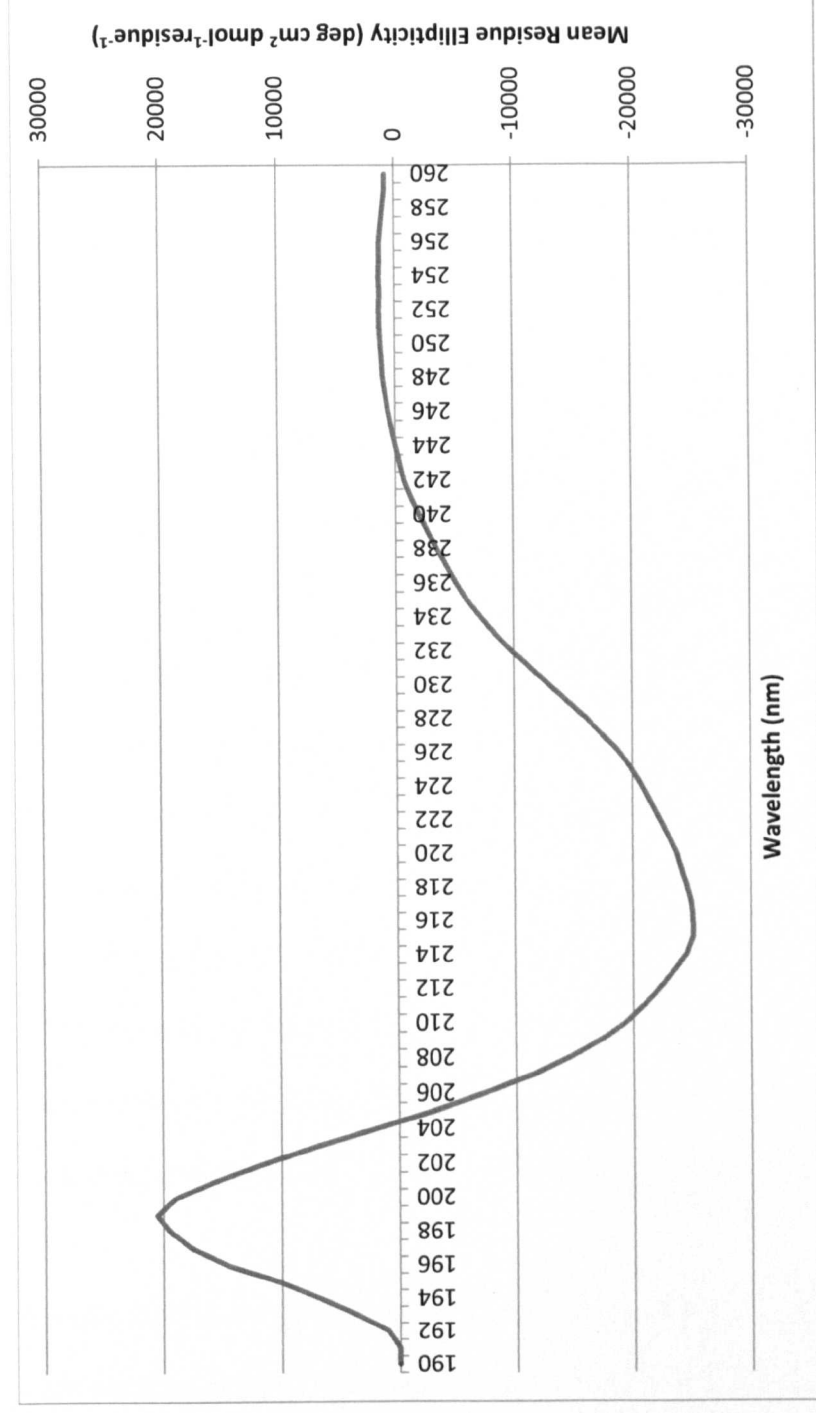


Figure 3.9 Far UV spectrum of OBP3 in 10 mM Potassium phosphate buffer, pH 7.0. The minimum between 205 -240 nm is indicative of high anti-parallel β -sheet content.

3.5.3.3 Attempted crystallisation of OBP3

In order to determine the high resolution structure of OBP3, x-ray crystallography was attempted. A number of lipocalins have previously been crystallised including nitrophorin, retinol binding protein and the human complement protein C8 γ (Amoia & Monfort 2007; Redondo et al. 2008; Kaufman & Sodetz 1994; Flower et al. 2000). Additionally the structures of bovine (Bianchet et al., 1996), porcine (Spinelli et al. 1998) and rat subtype 1 (White et al., 2009) odour binding proteins have been solved using x-ray crystallography (section 1.3).

Crystallisation of OBP3 was attempted using both sparse matrix and systematic screens of crystallisation agents at the University of Birmingham with the help of Dr. Scott White. Sparse matrix screens sample a large area of crystallisation space, by including a large range of precipitants, salt conditions, neutralized organic acids and solution pH values. Systematic screens at pH 7 and pH 8 were also explored as OBP3 had been purified successfully using these pHs. After continual monitoring of crystal trays stored at 18°C it was noted that a large number of drops contained amorphous precipitant, whilst a few remained clear, but failed to produce crystals or anything indicating crystals may form.

Initially OBP3 had been dialysed into pH 8 Tris buffer for crystallisation trials. An alternative buffer (10 mM KHPO₄, pH 7.0) was also used when the protein was sent to the laboratory of Prof. Sam Yong Park (Yokohama City University, Japan). Phosphate can lead to false-positive in crystallisation screens so buffer

only controls were required. Micro-crystals were detected in 0.1 M ammonium sulphate, 15 % polyethylene glycol, however this could not be optimised to produce crystals that diffracted.

3.2.5 Assessment of the thermal stability of OBP3

In order to more fully characterise OBP3 the stability of the protein was assessed using circular dichroism (CD) and differential scanning calorimetry (DSC). It was useful to find the thermal stability of OBP3 for a number of reasons; firstly to discover if the melting temperature was similar to that of other odour binding proteins, secondly to be able to compare wild type OBP3 with mutants of OBP3 and thirdly to gain an indication of how stable OBP3 was likely to be over the course of relatively long (in excess of a week) NMR experiments run during assignment data collection (see chapter 4).

3.2.5.1 Determining the melting temperature of OBP3 using circular dichroism

When a protein unfolds as it is thermally denatured a loss of protein secondary structure is seen. This can be monitored using CD, which, as described in section 3.2.4.2, gives a signal reflecting the protein secondary structure. Observing figure 3.14, a signal minimum at 217 nm is representative of the β -sheet component of OBP3 secondary structure. The CD signal at 217 nm was monitored as the temperature was raised in 1 K intervals from 298 K to 368 K. Figure 3.10 shows the melting curve obtained.

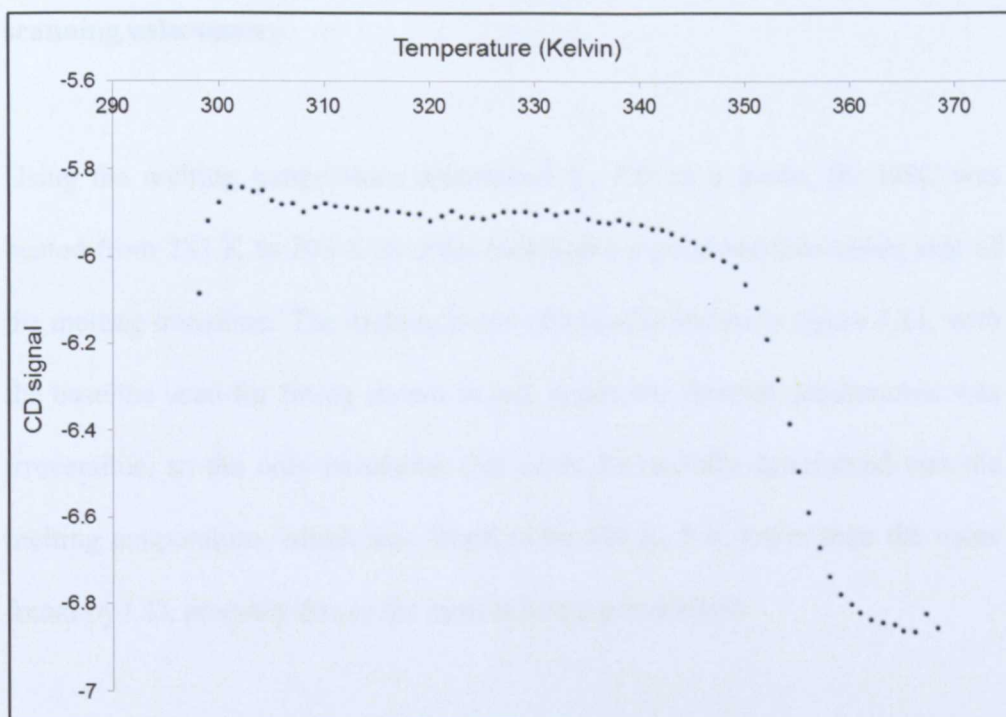


Figure 3.10 Melting curve of OBP3 monitored by circular dichroism, by fitting the curve the T_m was found to be 355 K.

The lack of a straight baseline at lower temperatures made fitting more difficult (described in section 2.6.5.1), therefore the first 3 data points were excluded from the baseline. A melting temperature of 355 K was established. The sample was cooled from 368 K back down to 298 K; however it was apparent that the protein had been irreversibly denatured. It was therefore not possible to determine the enthalpy of folding. Differential scanning calorimetry (DSC) was used as a complementary technique to confirm the melting temperature of OBP3.

3.2.5.2 Determining the melting temperature of OBP3 using differential scanning calorimetry

Using the melting temperature determined by CD as a guide, the DSC was heated from 283 K to 393 K in order to achieve a good baseline either side of the melting transition. The melting curve obtained is shown in figure 3.11, with the baseline used for fitting shown in red. Again the thermal denaturation was irreversible, so the only parameter that could be usefully determined was the melting temperature, which was found to be 346 K, 9 K lower than the value found by CD, possibly due to the system being pressurized.

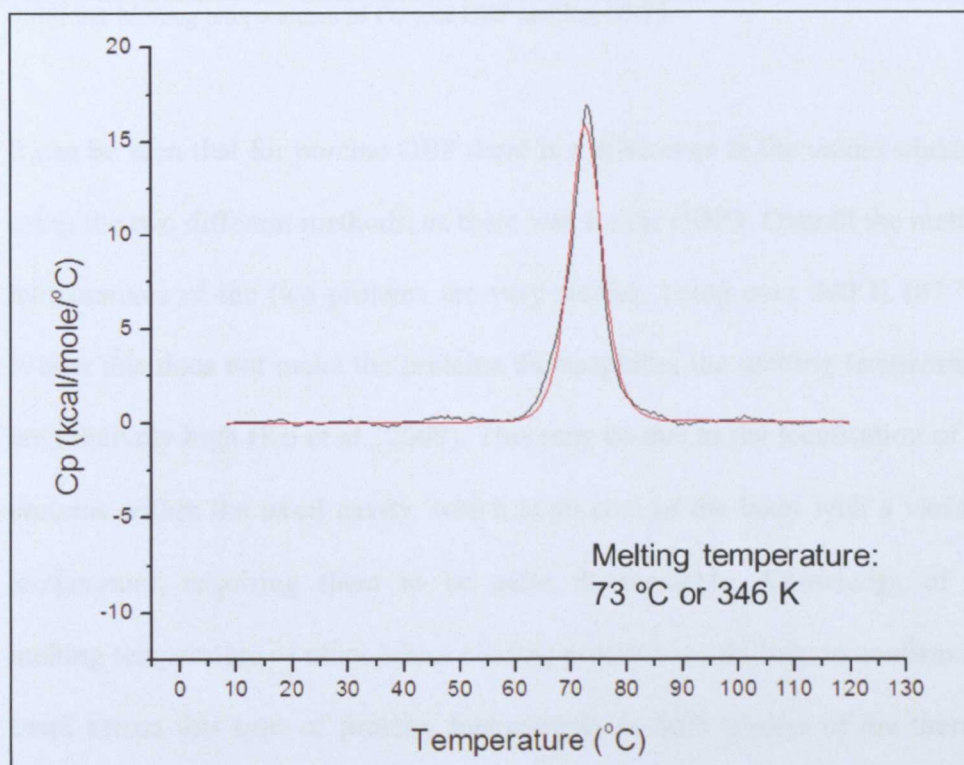


Figure 3.11 DSC melting curve of OBP3, heating the sample from 283 K to 293 K. A buffer to buffer baseline has been subtracted. The raw curve is shown in black, the red curve was used to fit the data and generate the melting temperature.

Of the other OBPs, porcine odour binding protein is the only protein with melting temperatures quoted in the literature. Table 3.1 compares the melting temperatures of porcine OBP and rat OBP3 using the methods of CD and DSC.

Odour Binding Protein	T _m (K)	pH	Method	Reference
Porcine	347	not stated	CD melt	(Parisi et al., 2005)
Porcine	342	6.6	DSC	(Burova et al., 1999)
Rat OBP3	355	7.0	CD Melt	This thesis
Rat OBP3	346	7.0	DSC	This thesis

Table 3.1 Melting temperatures of Porcine OBP and Rat OBP3.

It can be seen that for porcine OBP there is a difference in the values obtained using the two different methods, as there was for rat OBP3. Overall the melting temperatures of the two proteins are very similar, being over 340 K (67 °C). Whilst this does not make the proteins thermophilic, the melting temperatures are relatively high (Ku et al., 2009). This may be due to the localisation of the proteins within the nasal cavity, which is an area of the body with a variable temperature, requiring them to be quite thermostable. Knowledge of the melting temperature of other odour binding proteins would help to confirm this trend across this type of protein. Interestingly in both studies of the thermal stability of porcine OBP the unfolding reaction was found to be reversible, despite the protein being raised to a temperature of 373 K which is higher than rat OBP3 was raised to in the CD experiment.

3.3 Conclusions

This chapter has described the successful overexpression and purification of rat OBP3. The expression system used enabled large quantities of protein to be produced, which is ideal for biophysical studies. The histidine-tagged OBP3 was successfully purified using nickel affinity and gel filtration chromatography. The method used was adapted from Löbel et al. (2002) and resulted in a higher overall yield of pure protein (50 mg per litre of bacterial culture).

Circular dichroism showed the protein produced had the expected secondary structure; high β -sheet content with some α -helical content. Proton NMR suggested the protein was folded, as judged by the good dispersion of proton signals. Attempts to crystallise the protein in order to solve the high resolution structure of OBP3 were unsuccessful, however crystallisation in the future may still be possible. This may be achieved by first saturating the protein with a ligand, to avoid the possibility that crystal formation was being prevented by odour binding and unbinding. Alternatively more crystallisation conditions could be explored including; more buffer conditions, crystal tray storage temperature and protein concentration.

Finally, CD and DSC revealed that OBP3 was relatively thermally stable, with T_m values of 355 K and 346 K found respectively, which were comparable to porcine OBP. The thermal stability of these proteins may prove important for withstanding their extracellular environment, which is affected by nasal air flow that varies in temperature.

The OBP3 produced was suitable for biophysical binding studies using isothermal titration calorimetry (ITC) and nuclear magnetic resonance spectroscopy (NMR). For NMR studies of OBP3 the first stage, described in the following chapter, was to assign the protein.

4. NMR ASSIGNMENT OF OBP3

4.1 Introduction

ITC, the other technique used in this study to look at binding interactions of OBP3, provides a wealth of information about the binding characteristics of OBP3 (both affinity and thermodynamic parameters). However, the data provided are global in nature, reflecting changes in the state of the protein culminating from the sum of individual interactions and changes. Nuclear Magnetic Resonance (NMR) can be complementary to ITC data but provides resolution on a per residue basis. Although NMR analysis is more time consuming than other biophysical techniques the amount of information available on ligand binding affinities, exchange rates, the residues involved in ligand binding, measurement of protein dynamics and structure makes it worthwhile.

The OBP3 system is particularly suited for NMR study as large amounts of protein can be produced in a matter of days (50 mg of protein from 1 litre of bacterial culture, which was enough to make four 1 mM samples for NMR). This is also ideal for isotopically labelling the protein for heteronuclear NMR experiments, when growing bacteria in minimal media invariably decreases overall protein yield. Additionally NMR data from other vertebrate odour binding proteins has not been obtained and therefore OBP3 warranted study. OBP3 has high sequence identity to the Major Urinary Proteins (in excess of 80 %, see section 1.4) and such proteins have been studied by NMR. These proteins, like OBP3, are around 20 kDa in size and have proved to be a suitable

size for studying the whole protein by conventional NMR experiments.(Abbate et al. 1999; Lücke et al. 1999; Bingham et al. 2004)

The largest protein NMR assignment to date, of a single protein, was that of Malate synthase G, which contains 723 residues of which 654 of the backbone nitrogen atoms were assigned (Tugarinov et al., 2002; Grishaev et al., 2008). This assignment was completed using an 800 MHz spectrometer over 8 years ago, and higher field spectrometers (now up to 1 GHz) haven't as yet succeeded in surpassing this despite the predicted improvement in spectral resolution when using Transverse Relaxation Optimised Spectroscopy (TROSY) approaches at very high field (Pervushin et al. 1997; Lin et al. 2000).

The TROSY experiment suppresses transverse relaxation in multidimensional NMR experiments (such as the $^1\text{H} / ^{15}\text{N}$ HSQC described below). Interference between dipole-dipole coupling and chemical shift anisotropy (CSA), which mutually cancel, is used, giving a single sharp peak in the spectrum, rather than broad line widths, in decoupled HSQC experiments, typically characteristic of larger proteins. Generally high magnetic fields are required as CSA scales with field strength.

For protein-ligand binding studies the $^1\text{H} / ^{15}\text{N}$ Heteronuclear Single Quantum Coherence (HSQC) experiment is the key tool (Cavanagh et al., 2006). It is a two-dimensional experiment which correlates the ^1H and ^{15}N chemical shifts of amide groups, yielding a peak for each residue in the protein (except prolines which have secondary amide groups), plus peaks for any side chains with

nitrogen-bound protons (Arginine, Lysine, Glutamine and Asparagine). It requires nitrogen-15 as this has a nuclear spin of $\frac{1}{2}$, which is observable by NMR. The $^1\text{H} / ^{15}\text{N}$ experiment may in theory be conducted using the natural abundance of ^{15}N . However, NMR sensitivity for nitrogen is 1.04×10^3 times less than for proton, whilst the natural abundance of nitrogen is 0.37 %, meaning at natural abundance level the sensitivity for nitrogen is reduced to 3.8×10^{-6} that of proton (Gust et al., 1975) making it impossible to use at natural abundance for molecules at low concentrations.

For larger proteins (typically those over 10 kDa) there is considerable overlap in the $^1\text{H} / ^{15}\text{N}$ HSQC. There is also a decreased likelihood of finding suitable unique stretches of sequence to unambiguously assign residues. Three dimensional experiments such as the $^1\text{H} / ^{15}\text{N}$ HSQC-NOESY can be used; however, the presence of a large numbers of β -strands, as found in the odour binding proteins, complicates the analysis of the $^1\text{H} / ^{15}\text{N}$ HSQC-NOESY, due to cross-strand NOEs from short NH-NH distances.

Isotopically labelling with ^{13}C and ^{15}N allows the use of triple resonance three dimensional experiments that are designed to easily connect amino acid residues in sequence (Sattler, 1999). Figure 4.1 shows how these experiments are used, demonstrated with the $\text{C}\alpha$ and $\text{C}\beta$ shifts. In one experiment, such as a CBCANH experiment (shown in orange), the carbon shifts of a particular residue (*i*) are recorded at nitrogen frequency “X”, as well as the carbon shifts of the preceding residue (*i-1*, shown in purple), these peaks are usually weaker. In another experiment, (such as the CBCA(CO)NH), only the shifts of the

preceding residue are recorded (in purple, with a dotted orange line) at the same nitrogen frequency. The actual nitrogen frequency (Y) of these peaks is then found, and movement is made backwards along the sequence. (This can also be done moving forward along the protein backbone).

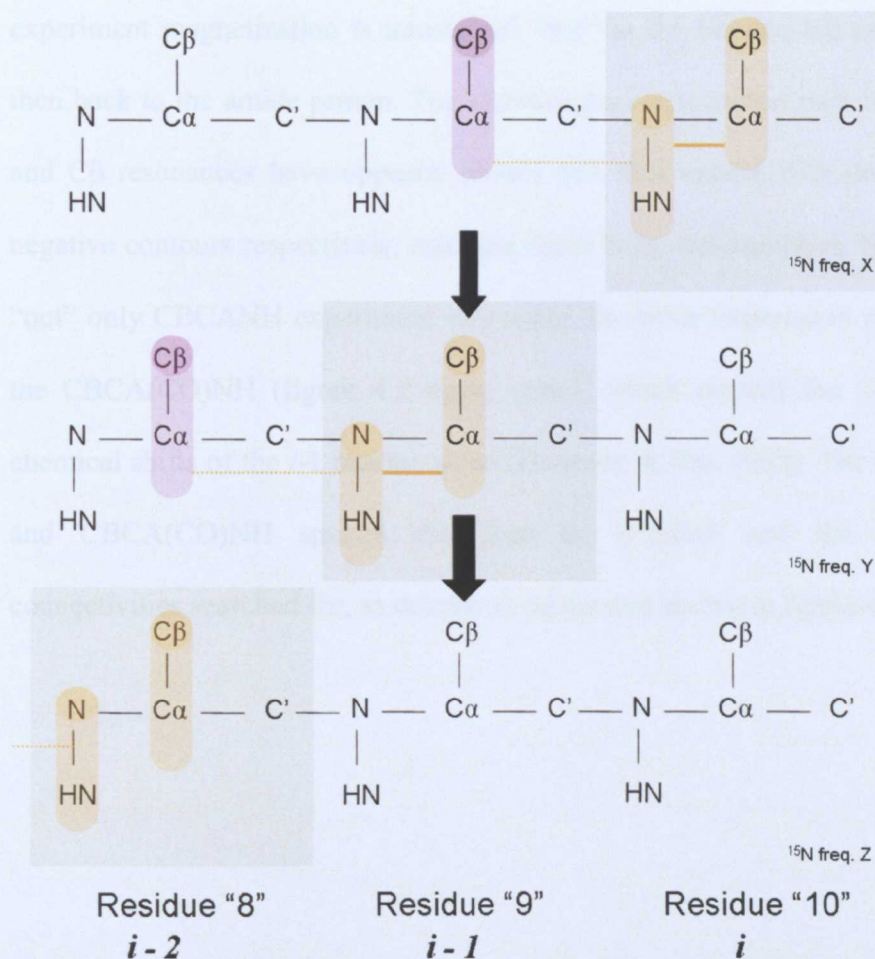


Figure 4.1 Sequential assignment using triple resonance experiments. Described more fully in the text, assignment is made from residue 10 (highlighted in grey at the top right) and connectivity made to residue 9 and then residue 8, moving backwards along the chain. The i experiment (such as the CBCANH), records the atoms shown in orange and purple, whilst the $i-1$ experiment (such as the CBCA(CO)NH) records just those atoms shown in purple.

To determine $C\alpha$ and $C\beta$ chemical shifts two experiments are recorded; the first records the $C\alpha$ and $C\beta$ chemical shifts of residue i and residue $i-1$. There are different experiments for collecting this information; the CBCANH (Grzesiek & Bax 1992a) and the HNCACB (Wittkind & Mueller 1993). The transfer of magnetization is shown in figure 4.2. In the case of the HNCACB experiment magnetization is transferred “out” to the $H\alpha$ and $H\beta$ protons and then back to the amide proton. The experiments are recorded such that the $C\alpha$ and $C\beta$ resonances have opposite phases and thus appear with positive and negative contours respectively, enabling them to be distinguished. Initially the “out” only CBCANH experiment was used. The other experiment recorded is the CBCA(CO)NH (figure 4.2 upper pane), which records the $C\alpha$ and $C\beta$ chemical shifts of the $i-1$ residue alone (Grzesiek & Bax 1992). The CBCANH and CBCA(CO)NH spectra can then be overlaid and the sequential connectivities searched for, as described earlier and shown in figure 4.1.

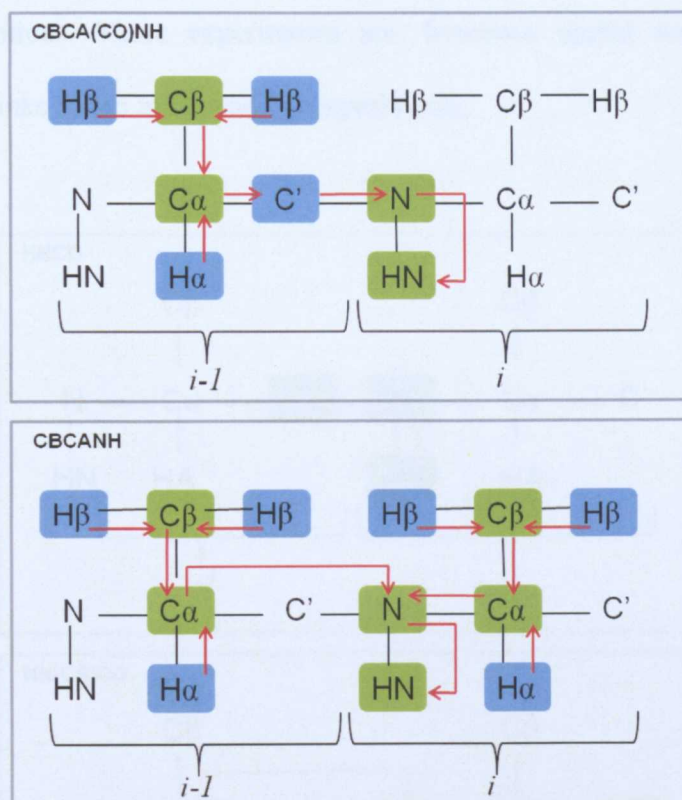


Figure 4.2 Transfer of magnetization in the CBCA(CO)NH (top panel) and CBCANH (bottom panel) experiments. The flow of magnetization is shown by the red arrows. Blue boxed atoms indicate magnetization flows through them whilst green boxed atoms are observed. In both experiments the amide proton is used for detection.

To determine the carbonyl shifts the HN(CA)CO (Clubb et al. 1992) and HNCO (Kay et al. 1990; Grzesiek & Bax 1992b; Muhandiram & Kay 1994) experiments are used together, in a similar manner to the $C\alpha$ and $C\beta$ shift experiments. The magnetization flow diagrams for these experiments are shown in figure 4.3. The HN(CA)CO has two peaks, one for residue *i* and one for residue *i-1*, whilst the HNCO contains only the *i-1* peaks. Finding matching peaks is more complicated than for $C\alpha$ and $C\beta$ shifts as there is a lot of overlap, due to a smaller frequency range in the carbon dimension. Additionally only one peak is being searched for, so it is harder to be certain when there are

multiple options. These experiments are, however, useful for confirming sequential links found in $C\alpha$ and $C\beta$ experiments.

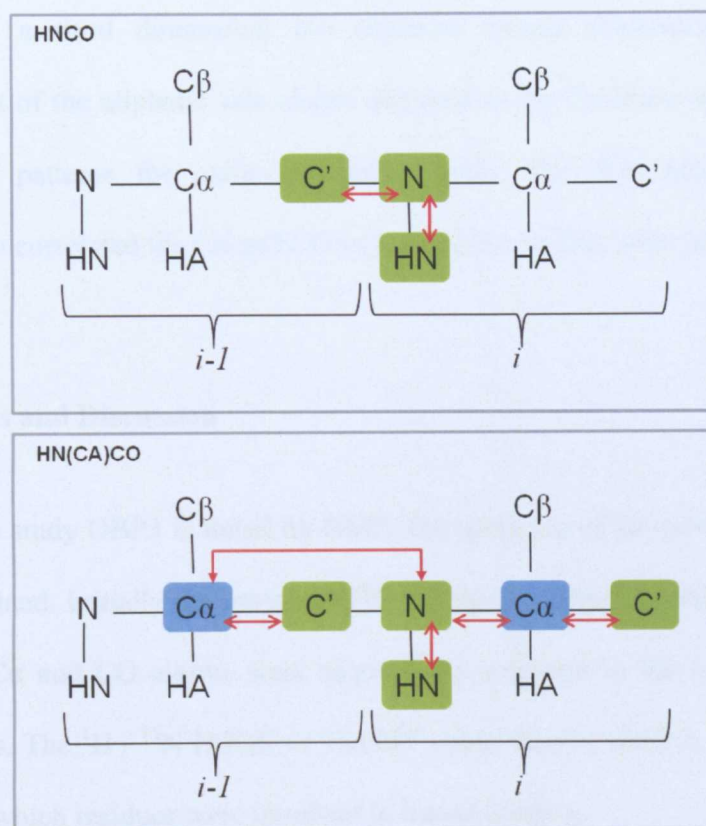


Figure 4.3 Transfer of magnetization in the HNCO (top panel) and HN(CA)CO (bottom panel) experiments. Colouring is the same as for figure 4.2. In the case of the HNCO experiment magnetization is transferred from the amide proton to the nitrogen and then selectively to the carbonyl carbon. The magnetization is passed back to the amide proton for detection. The HNCO is a very sensitive experiment. In the HN(CA)CO experiment magnetization flows from the nitrogen to the $C\alpha$ and then to the carbonyl and back to the amide proton for detection.

In addition direct carbon observation experiments were used. This means protons aren't used for detection and therefore aqueous buffers don't introduce solvent artefacts and the spectra are clearer. The CACO experiment correlates the $C\alpha$ and CO shifts and thus provides confirmation that particular carbon

shifts (from experiments such as the HNCACB and HNCACO) belong to the same residue. This is useful when there is overlap in the proton dimension at a particular nitrogen frequency. The CCCO-TOCSY (Bermel et al. 2005, 2006) then adds a third dimension, the aliphatic carbon dimension; allowing assignment of the aliphatic side chains attached to the C α atoms and provides distinctive patterns for particular amino acids. The HACACO is a 3D experiment correlated the C α and CO of a particular residue with its H α atom.

4.2 Results and Discussion

In order to study OBP3 in detail by NMR, the spectrum of the protein first had to be assigned. Initially the protein backbone signals, which consist of the H_N, N_H, H α , C α and CO atoms, were sequentially assigned to the correct NMR resonances. The ¹H / ¹⁵N HSQC or TROSY could then be used to monitor, for example, which residues were involved in ligand binding.

4.2.1 Homonuclear experiments

A 1 mM sample of non-isotopically labelled OBP3, in KHPO₄ at pH 7 was used to run the one-dimensional proton Watergate experiment shown in section 3.2.4. This confirmed the protein was folded and free from contaminants and artefacts, prior to longer experiments which required labelled protein. This sample was also used to run 2D NOESY and TOCSY experiments. The peaks are very overlapped, as would be expected from a 19 kDa protein, due to the high number of protons, and hence the spectra are very difficult to use. It is

also likely that the signal from solvent water has obscured some of the H α protons. To improve these spectra and remove the effect from solvent water the same experiments were run on a D₂O (²H₂O) exchanged sample. This required dissolving lyophilised OBP3 in D₂O, lyophilising the sample again and then redissolving the sample again. This was repeated five times to try to exchange the labile protons in OBP3 with deuterium, leading them to no longer be visible in the NMR spectrum (Wüthrich & Wagner 1979, 1982; Jue et al. 1984). A full description of how this sample was prepared is described in section 2.6.3.2. The resulting TOCSY spectrum is shown in figure 4.4. It is interesting to note that despite numerous resuspension and freeze drying rounds there remains a number of unexchanged protons in the amide region (shaded grey area). On their own these spectra were not very useful for assigning OBP3 but were used to heteronuclear experiments to aid the deconvolution of the proton dimension for side chain identification and assignment.

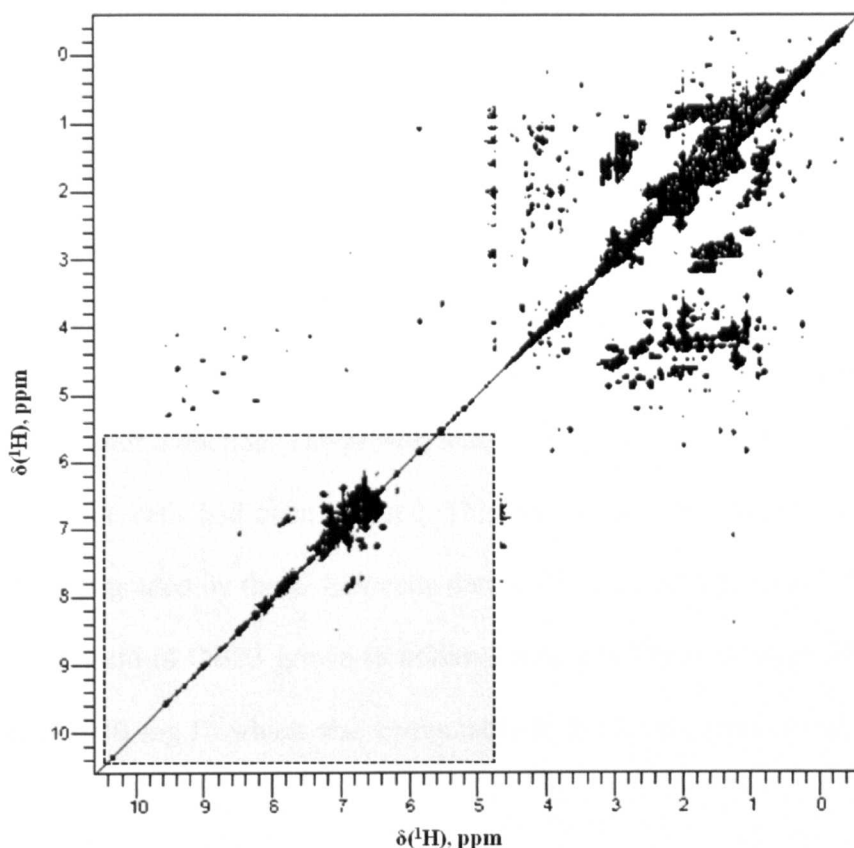


Figure 4.4 TOCSY spectrum of D₂O exchanged OBP3. Due to the number of residues considerable overlap is seen. A number of TOCSY cross peaks to H α protons are still present in the amide region, indicating unexchanged protons. These could be due to highly protected residues lining the putative binding pocket.

4.2.2 Isotopic labelling

Isotopically labelled protein was produced in a minimal medium. A sole source of the required isotope, for protein NMR nitrogen-15 and carbon-13 is used. To produce isotopically labelled OBP3 using *E. coli* ¹⁴N ammonium chloride was replaced with ¹⁵N ammonium chloride, and ¹²C glucose was replaced with ¹³C₆ glucose as the sole carbon source. The minimal media used is described in section 2.2.5.

An initial trial growth of M15 cells containing pQE31 plasmid in the minimal medium indicated that growth was slower. The cells took 4 hours (rather than 1-2 hours) to reach the OD_{600nm} of 0.6 required prior to induction of protein overexpression and the protein yield after eight hours was less than 10 mg l^{-1} (rather than 50 mg l^{-1} achieved using LB media and the procedure described in section 2.2.1). The growth regime was therefore altered to allow a 16 hour growth period after induction. The growth temperature was dropped from 37°C to 25°C after the cells had been induced. This was to limit the chance of the protein being degraded by the *E. coli* cells during the long growth period. This increased the yield of OBP3 grown in minimal media (without isotopic labels being used) to 40 mg l^{-1} which was comparable to the levels achieved in LB media.

Purification was carried out using the same method described in section 2.3. Figure 4.5 shows pure $^{15}\text{N} / ^{13}\text{C}$ labelled OBP3 protein by SDS-PAGE after gel filtration. $20\text{ }\mu\text{l}$ from each 5 ml fraction which contained protein was loaded onto the SDS-PAGE gel. The thick band reflects the high yield of pure protein was achieved. After desalting chromatography and lyophilisation (as described in section 2.3.4) a final yield of 25 mg l^{-1} was achieved. When producing ^{15}N OBP3 yields of greater than 35 mg l^{-1} were routinely achieved.

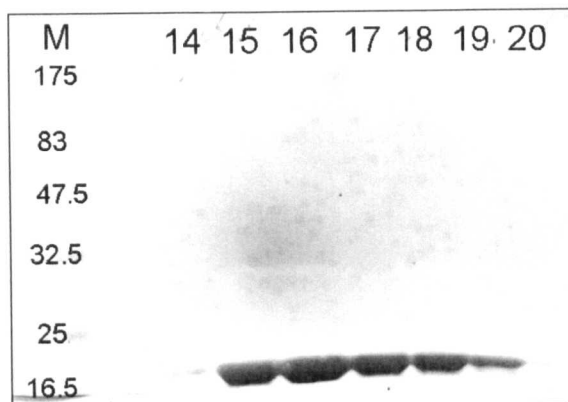


Figure 4.5 SDS PAGE of gel filtration fractions (shown along the top) of ^{15}N OBP3 purification. The protein size marker (M) indicated an apparent molecular weight of approximately 18-19 kDa, corresponding to OBP3

4.2.3 Heteronuclear Experiments

The first heteronuclear experiment recorded was a $^1\text{H} / ^{15}\text{N}$ TROSY (Pervushin et al., 1997) at 298K in pH 7.0 potassium phosphate buffer on a 1mM sample of ^{15}N OBP3 (described in section 2.2.5). The spectrum is shown in figure 4.6. It shows the peaks are well dispersed, which demonstrates that all the amide protons are in distinct chemical environments. This indicates the protein is folded and contains no obviously unstructured regions which would result in a collapsed spectrum with all the peaks appearing in one area (Chasman, 2003).

OBP3 consists of 160 residues, plus 13 additional residues as part of the histidine tag, which may not be visible on the spectrum due to it exchanging on a different timescale. An initial peak count found 155 peaks, which allowing for overlap of some peaks suggests almost all residues of the main chain of OBP3 were visible.

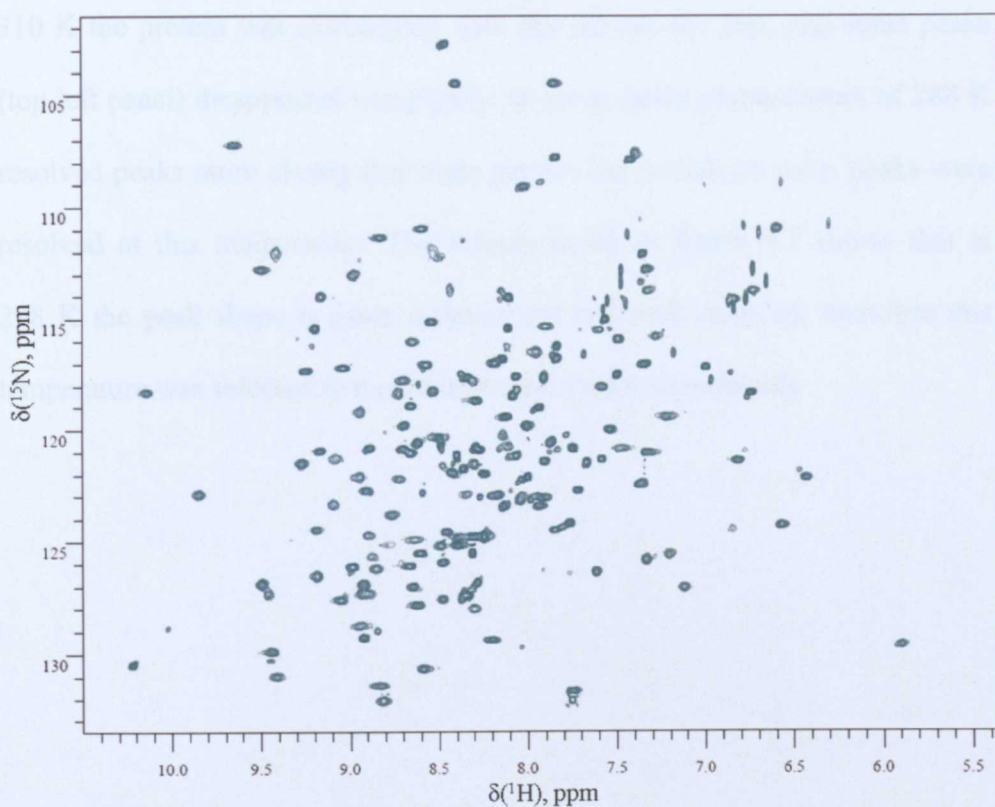


Figure 4.6 A $^1\text{H} / ^{15}\text{N}$ TROSY of OBP3 in pH 7.0 10 mM potassium phosphate buffer, recorded at 298K. The peaks are well dispersed and indicated assignment of OBP3 would be possible

In order to optimise the number of peaks prior to triple resonance experiments being run a decrease in pH was tried, as well as varying the sample temperature. Lowering the pH to 6.1 made no difference to the number of discernable peaks; however a slight movement of peak positions appeared to resolve some peaks more clearly, so this pH was selected for triple resonance experiments.

The $^1\text{H} / ^{15}\text{N}$ TROSY was also recorded at a lower (288 K) and higher (310 K) temperature, with the aim of revealing additional peaks. Figure 4.7 shows an

overlay of the spectra at the three different temperatures. It is apparent that at 310 K the protein was exchanging with the solvent too fast, and some peaks (top left panel) disappeared completely. In some cases a temperature of 288 K resolved peaks more clearly (top right panel), but overall no more peaks were resolved at this temperature. The bottom panel in figure 4.7 shows that at 298 K the peak shape is quite symmetrical and well resolved; therefore this temperature was selected to record triple resonance experiments.

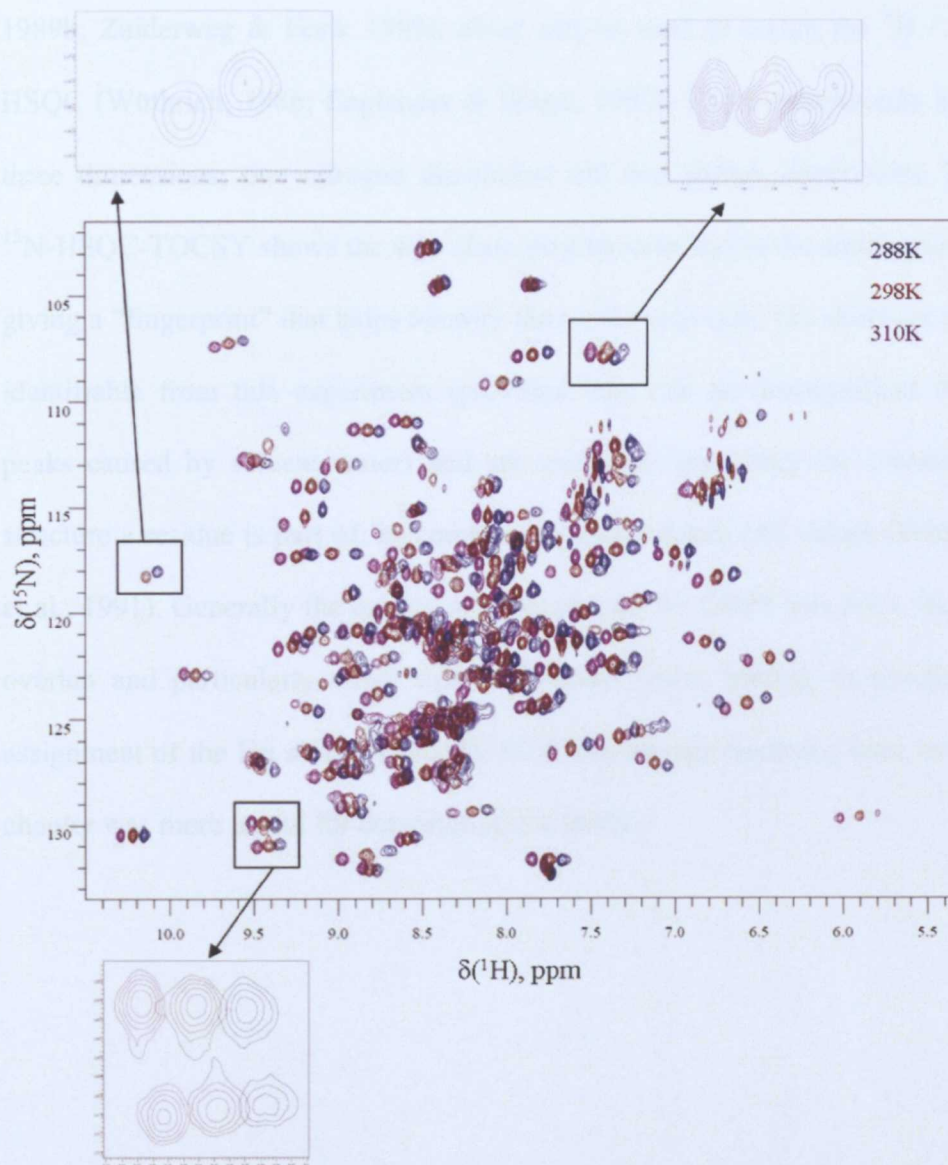


Figure 4.7 $^1\text{H} / ^{15}\text{N}$ TROSY spectra of OBP3 at different temperatures. Peaks are shifted at different temperatures. A large reduction in the number of peaks was seen at 310 K. Some peaks are better resolved at 288 K, however overall most peaks were found at 298 K.

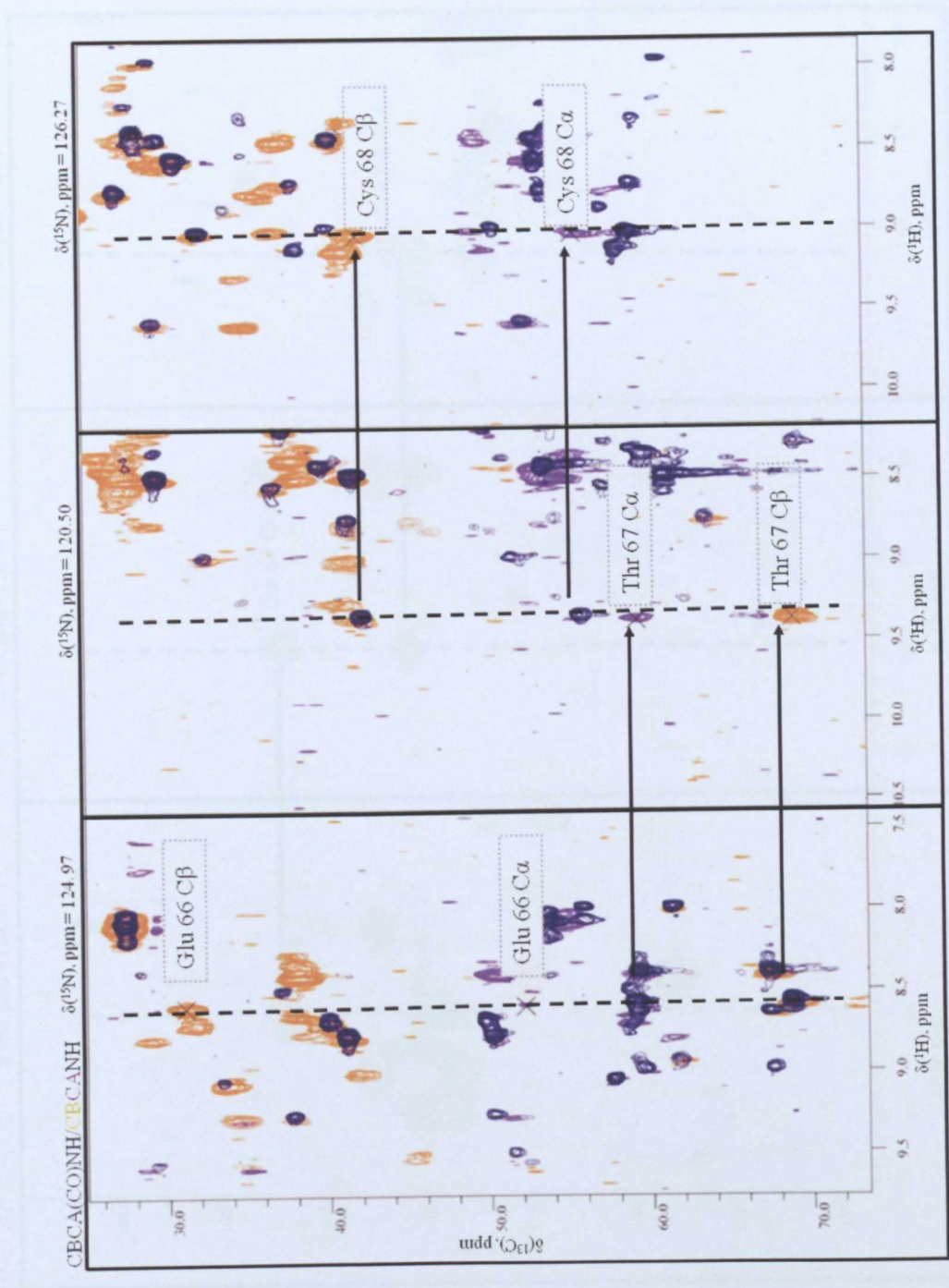
4.2.4 Three Dimensional $^1\text{H} / ^{15}\text{N}$ Heteronuclear NMR experiments

As described in the introduction to this chapter, for small proteins, double resonance experiments, such as the $^1\text{H} / ^{15}\text{N}$ -HSQC-TOCSY (Wijmenga et al. 1989; Marion et al. 1989) and $^1\text{H} / ^{15}\text{N}$ -HSQC-NOESY (Marion et al. 1989a,

1989b; Zuiderweg & Fesik 1989), alone can be used to assign the ^1H / ^{15}N HSQC (Wüthrich, 1986; Englander & Wand, 1987). These experiments have three dimensions; one nitrogen dimension and two proton dimensions. The ^{15}N -HSQC-TOCSY shows the side chain protons attached to the amide proton, giving a “fingerprint” that helps identify the amino acid type. $\text{H}\alpha$ shifts are also identifiable from this experiment (provided they can be distinguished from peaks caused by solvent water) and are useful for predicting the secondary structure a residue is part of, by comparison with random coil values (Wishart et al., 1991). Generally the quality of these spectra for OBP3 was poor, due to overlap and particularly weak signals in some cases, leading to unreliable assignment of the $\text{H}\alpha$ shifts. The HACACO experiment discussed later in this chapter was more useful for determining $\text{H}\alpha$ shifts.

4.2.5 Triple Resonance NMR experiments

Triple resonance experiments were run on a 1 mM OBP3 sample isotopically labelled with ^{13}C and ^{15}N . Examples of the spectra collected are shown in figure 4.8. It was noted that the CBCANH experiment was particularly low in sensitivity with numerous scans (and therefore a long collection period, in excess of 5 days) required to achieve spectral quality that was good enough to use for assignment. The $\text{C}\alpha$ shifts were missing for a large number of residues. To overcome this a HNCA spectrum was acquired (Kay et al. 1990), this was quicker to collect (~24 h) and had markedly improved signal to noise and therefore the majority of $\text{C}\alpha$ resonances could be easily identified, though not assigned (>90%).



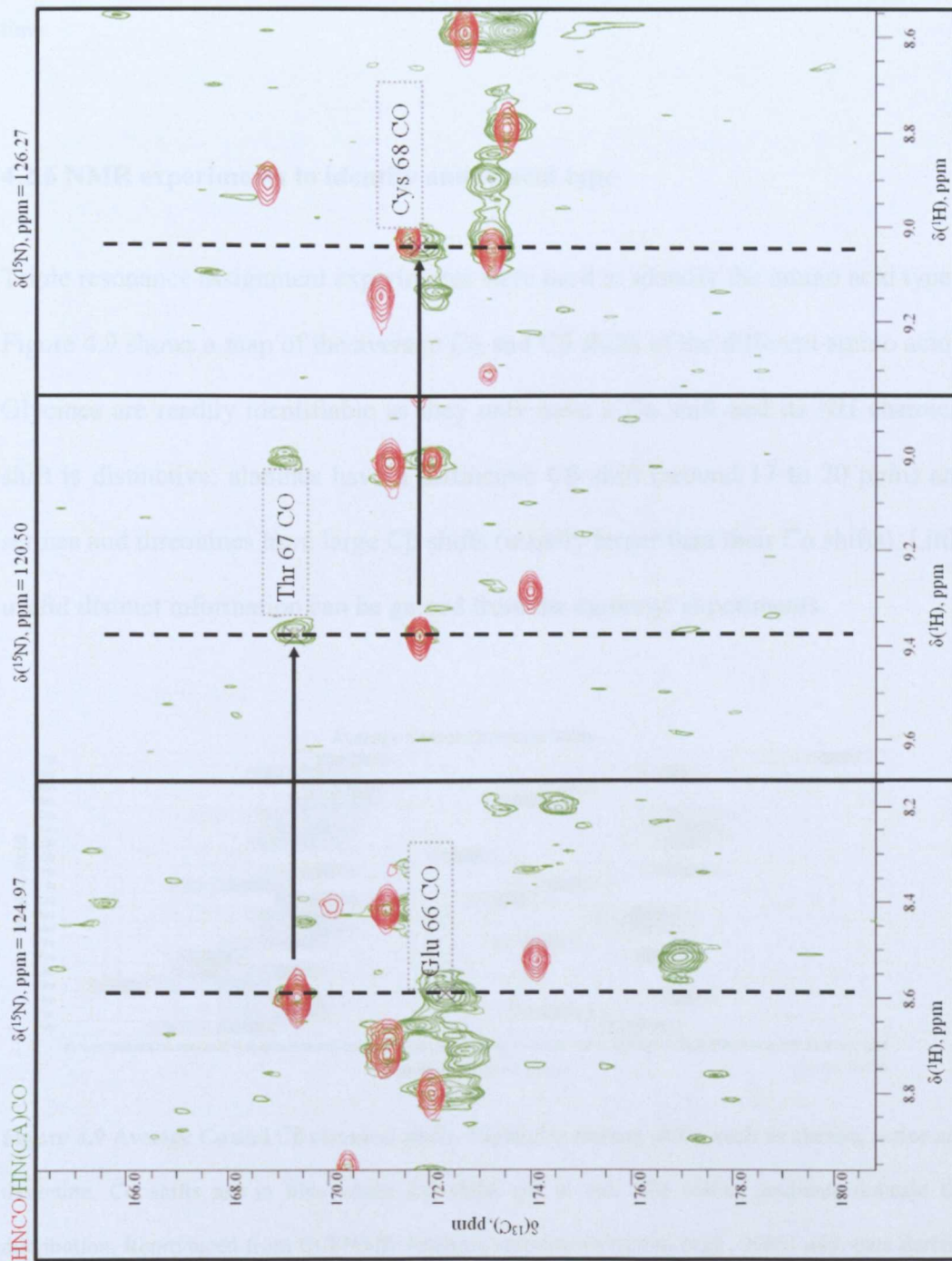


Figure 4.8 Examples of triple resonance experiments collected on OBP3. The experiment title colours correspond to the colours in the spectra. The HNC0 is clearly the most sensitive experiment as was expected. Generally the other experiments were high enough quality to use for assignment, however the CBCANH experiment was particularly insensitive and required a long period of experimental time.

4.2.6 NMR experiments to identify amino acid type

Triple resonance assignment experiments were used to identify the amino acid types. Figure 4.9 shows a map of the average $C\alpha$ and $C\beta$ shifts of the different amino acids. Glycines are readily identifiable as they only have a $C\alpha$ shift and its NH chemical shift is distinctive; alanines have a distinctive $C\beta$ shift (around 17 to 20 ppm) and serines and threonines have large $C\beta$ shifts (usually larger than their $C\alpha$ shifts). Little useful distinct information can be gained from the carbonyl experiments.

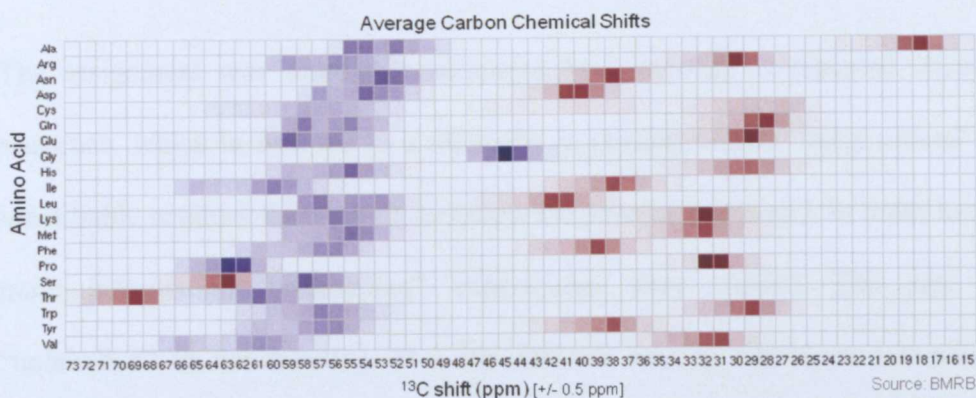


Figure 4.9 Average $C\alpha$ and $C\beta$ chemical shifts. Highlights distinct shifts, such as alanine, serine and threonine. $C\alpha$ shifts are in blue whilst $C\beta$ shifts are in red. The colour gradients indicate the distribution. Reproduced from CCPNMR Analysis software (Vranken et al., 2005) with data derived from the Biological Magnetic Resonance Bank.

The primary sequence of OBP3 contains the distinctive pair of amino acids Ala – Ser. In the CBCANH experiment the C β of Ser has a large shift (62 – 66 ppm), which is usually bigger than the C α shift and in the CBCA(CO)NH experiment at the same proton and nitrogen frequency the *i-I* experiment has a C β shift of 17 – 20 ppm) corresponding to the preceding alanine. This occurs twice in the primary sequence (Ala 3 – Ser 4 and Ala 25 – Ser 26) so enabled identification of these amino acids and narrowed down the likely position of these resonances to two places in the sequence. Ser and Thr are difficult to distinguish so therefore these resonances could also be Ala – Thr, however this sequence is not present in OBP3 so could be ruled out. Other distinctive sets of residues in OBP3 included; Gly 36 – Ser 37, Thr 89 – Phe 90 – Thr 91, Ser 127 – Ser 128, Thr 152 – Lys 153 – Thr 154. At this point about 30% of the residues could be assigned with near certainty.

4.2.6.1 “Unlabelling” experiments

The assignment was somewhat hampered by a number of repeated elements of sequence, which in turn led to a difficulty in placing certain “entry points” (easily identifiable residues from which to follow the protein backbone). In order to provide more information “unlabelling” experiments were trialled. The aim of the “unlabelled” or non-labelled ^1H / ^{15}N TROSY was to add more information on residue type. Conventionally single residue ^{15}N labelling approaches have been used (Muchmore et al. 1989; Goto & Kay 2000; Tanio et al. 2009); however this is a relatively expensive approach. Additionally, matching spectra where relatively few signals exist is more challenging than identifying missing signals. Selecting to not isotopically label certain residues or “unlabelling” is an alternative approach, where by ^{14}N amino acids are supplemented to a bacterial minimal medium containing a

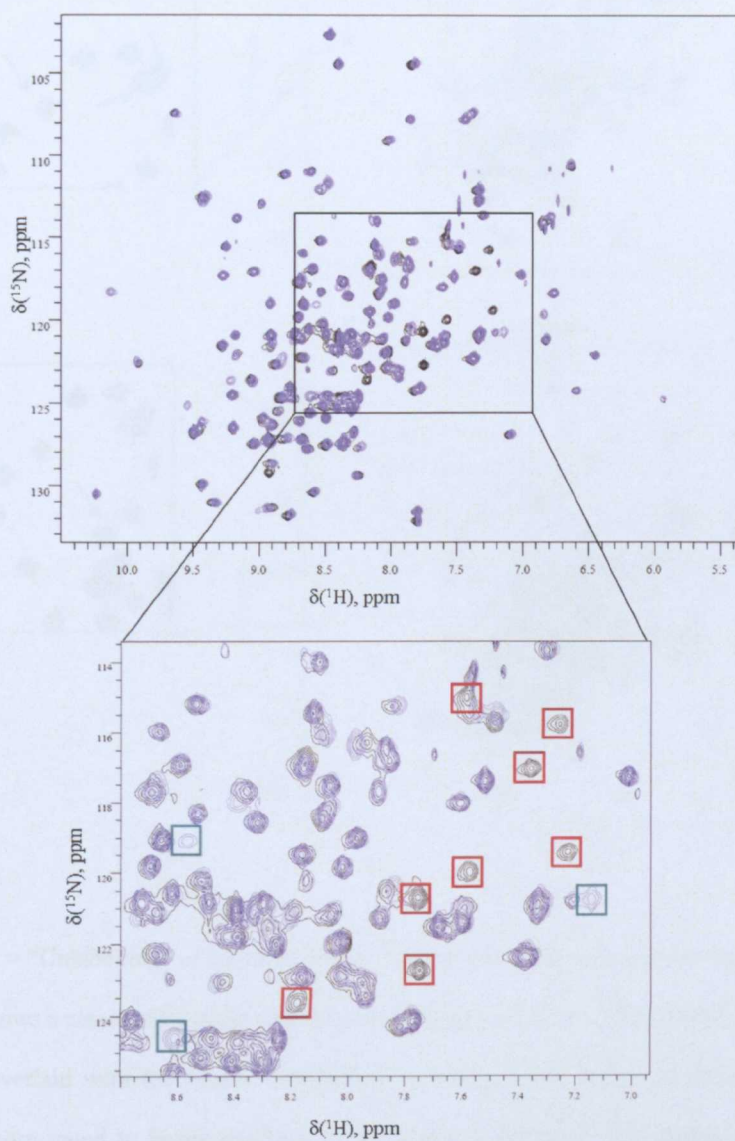
^{15}N nitrogen source (Atreya & Chary 2000; Krishnarjuna et al. 2011). The same procedure for ^{15}N isotopic labelling was used as described in section 2.2.2, except at the point of transfer to minimal media 0.5 g/l of the ^{14}N amino acid to be unlabelled was added, followed by another 0.5 g/l at the point of induction (described in section 2.2.6). Waugh's description of the amino acid production pathways in *E. coli* was used to choose amino acids that were terminal products and therefore not interconverted (Waugh, 1996).

The first amino acid trialled was lysine; this was because lysine is the most abundant amino acid in OBP3 (12 lysines out of 173 residues in total) that is a terminal product in the *E. coli* amino-acid biosynthetic pathways. The chance of finding peaks that had been reduced or had disappeared completely was therefore greatest. Figure 4.10 (A) shows the resulting spectrum; in red boxes are residues that are missing from the "unlabelled" spectrum. Those unlabelled all proved to be lysine residues (in total 11 of the 12 lysines expected were identified). In green boxes are some peaks that appear to be additional in the lysine "unlabelled" spectrum. The reason for this is unclear, but could be due a small amount of degraded protein.

In addition to lysine the technique was also attempted with a number of other amino acids. It was found that arginine and threonine "unlabelling" worked well (Figure 4.10 (B)) as did glutamine (however OBP3 only has 3 residues of this type). Phenylalanine and tyrosine proved to be more problematic as it appears these amino acids were interconverted and so the spectra collected were not interpretable. When attempting to unlabel cysteine, methionine and isoleucine it was found that the *E. coli* was unable to grow in media supplemented with these amino acids as they were

apparently toxic. (Harris, 1981; Karkhanis et al., 2007). Using unlabelling experiments gave more entry points and over 85% of the protein could be assigned.

A



B

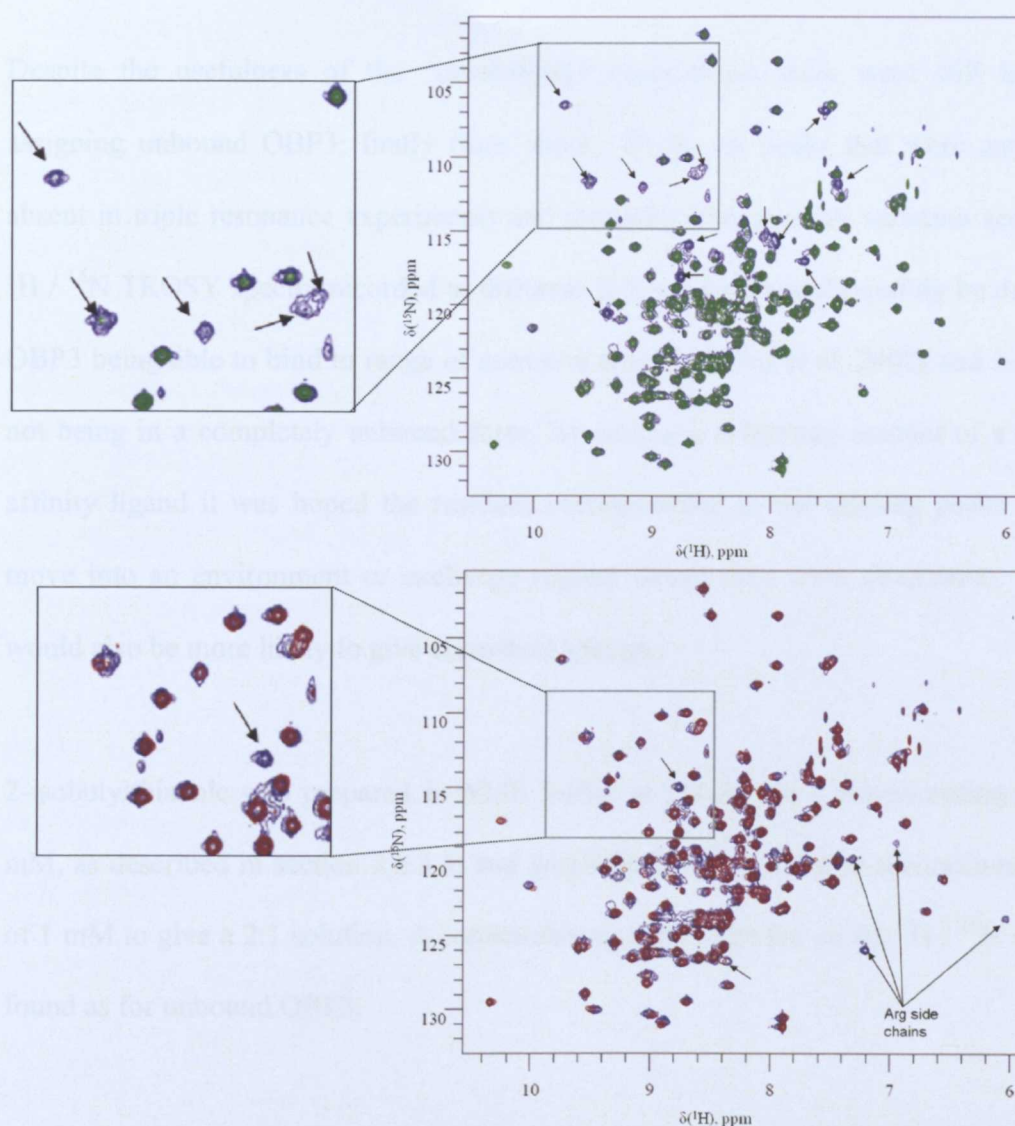


Figure 4.10 A = “Unlabelling” of lysine in OBP3. The top panel shows the entire spectrum, whilst the lower panel shows a closer view of the region most abundant in lysines. The original spectrum (shown in black) is overlaid with the lysine “unlabelled” spectrum (light blue). Boxed in red are peaks suggested to correspond to lysine residues. In green boxes are peaks that appear to be additional, possibly due to excess lysine carried over through protein production. B = “Unlabelling” of threonine top and arginine bottom in OBP3. The original spectra are shown in navy blue. The threonine “unlabelled” spectrum is shown in green and the arginine spectrum is shown in claret. Arrows point to the residues that have been “unlabelled”. In the arginine unlabelled spectrum the side chains are also unlabelled and are indicated on the spectrum. An area of each spectrum is enlarged to more clearly demonstrate the unlabelling.

4.2.7 Assignment of OBP3 bound to 2-isobutylthiazole

Despite the usefulness of the “unlabelling” experiments there were still issues assigning unbound OBP3; firstly from about 20 % of peaks that were entirely absent in triple resonance experiments and secondly from a small variation seen in $^1\text{H} / ^{15}\text{N}$ TROSY spectra recorded at different times, which could possibly be due to OBP3 being able to bind to range of common odours (Löbel et al. 2002) and in fact not being in a completely unbound form. By adding a saturating amount of a high affinity ligand it was hoped the residues corresponding to the missing peaks may move into an environment or exchange regime where they were observable. This would also be more likely to give consistent spectra.

2-isobutylthiazole was prepared in NMR buffer at pH 6.1, at a concentration of 2 mM, as described in section 2.6.3.5, and lyophilised OBP3 added to a concentration of 1 mM to give a 2:1 solution. A comparable number of peaks on the $^1\text{H} / ^{15}\text{N}$ were found as for unbound OBP3.

A similar set of triple resonance experiments were run to those described in section 4.2.5. Notably a HNCACB experiment was acquired to determine the $\text{C}\alpha$ and $\text{C}\beta$ chemical shifts and proved to give much better signal to noise and clearer $\text{C}\alpha$ chemical shifts than the CBCANH experiment run on unbound OBP3. This enabled over 60% of the protein to be assigned (with minimal reference to the unbound assignment which allowed a cross check of the initial assignment)

In addition, direct carbon detection experiments were also run including the CACO and HACACO experiments (Bermel et al. 2005, 2009). The HACACO spectrum was

extremely clear and the peaks were quite isolated. Due to this experiment using direct carbon observation the spectrum was not complicated by artefacts from solvent water suffered in the $^1\text{H} / ^{15}\text{N}$ HSQC TOCSY and NOESY. A full count of H_α chemical shift was visible and over 90% were assigned with the aid of the other direct carbon experiments.

The CACO spectra gives a protein “fingerprint” in a similar way to the $^1\text{H} / ^{15}\text{N}$ HSQC, and the assigned spectrum is shown in figure 4.11. There is a single peak for every amino acid, including proline residues, of which histidine-tagged OBP3 has one. A protein sample labelled only with carbon-13 was prepared as described in section 2.2.5 in order to run these experiments, these experiments could also have been run on protein labelled with both nitrogen-15 and carbon-13. The CACO experiment enabled confirmation of the previous assignments but also gave leads into another 10-15% of the assignments.

The assignment of 2-isobutylthiazole bound OBP3 was more straightforward than for unbound OBP3. The assignments were used in a number of cases to aid the unbound OBP3 assignment as the $^1\text{H} / ^{15}\text{N}$ TROSY spectra were quite similar so peaks were often easy to trace. The near complete assignments of the $^1\text{H} / ^{15}\text{N}$ TROSY spectra of unbound and 2-isobutylthiazole bound OBP3 are shown in figures 4.12 and 4.13 respectively. The shift values of the backbone amino acids are given in Appendix V and VI. Unfortunately most of β -strand G (particularly Met 102, Phe 103, His 104, Leu 105, Val 106, Asn 107 and Val 108) proved impossible to assign due to weak signals, meaning the residues in the strand could not be connected unambiguously. For the unbound protein assignment all the assignment were verified by comparison with the 2-isobutylthiazole bound assignment and in addition the $^1\text{H} / ^{15}\text{N}$ HSQC-NOESY was checked for the expected $\text{H}\alpha$ cross peaks. For the bound assignment the all assignments were confirmed using the CACO and CCCO-TOCSY and in appendix VI there are some residues with just $\text{C}\alpha$ or CO shifts, the values of which only originate from the direct carbon experiments and HNCA, HNCO or CBCA(CO)NH $i-1$ peaks.

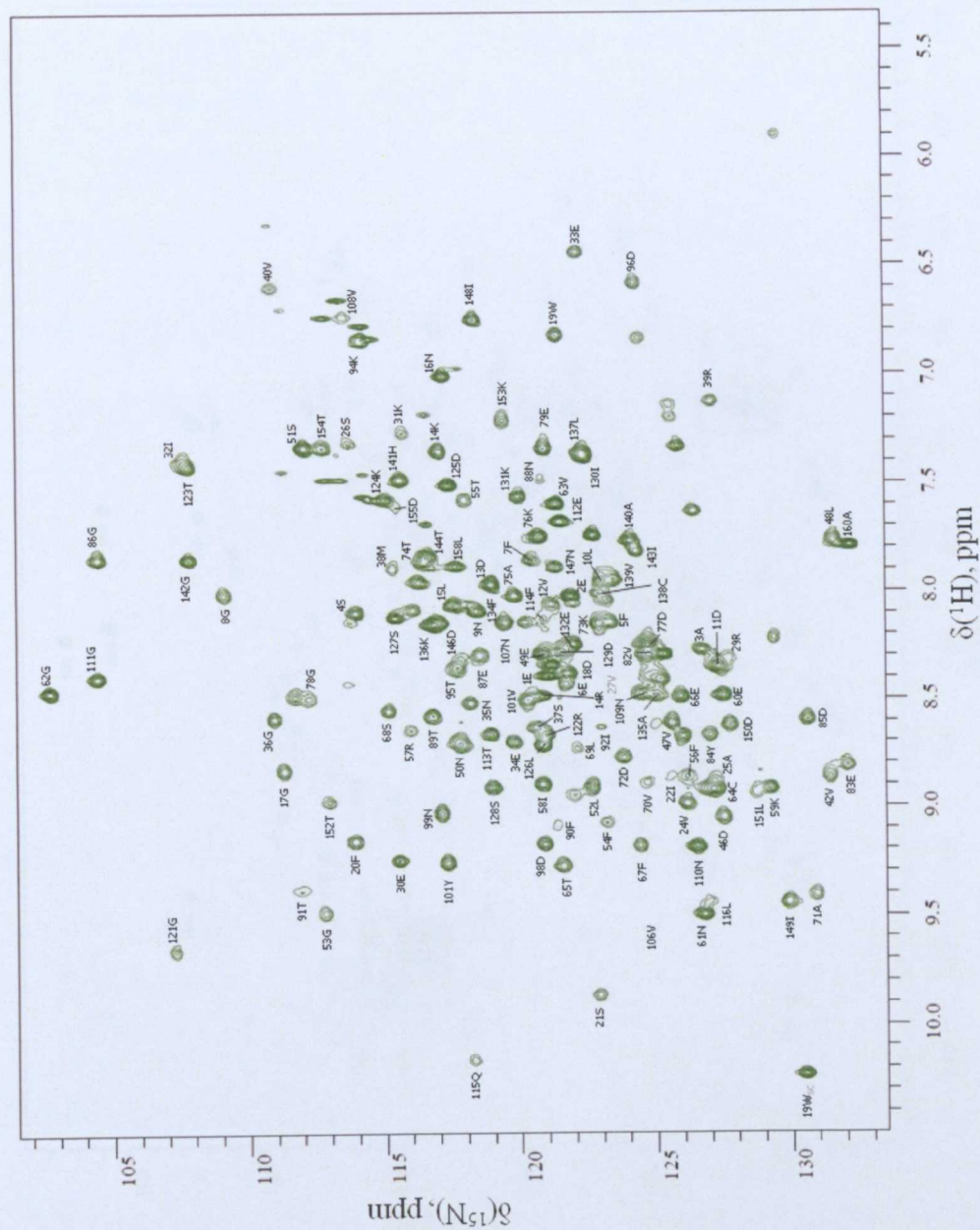


Fig 4.12 Assigned $^1\text{H} / ^{15}\text{N}$ TROSY spectrum of unbound OBP3 (green). The TROSY experiments means the NH side chains are not visible, in some cases they appear as weak peaks. In total 137 out of 160 residues have been assigned, plus 3 residues belonging to the histidine tag.

Fig 4.13 Assigned $^1\text{H} / ^{15}\text{N}$ TROSY spectrum of OBP3 bound to 2-isobutylthiazole (blue). In total 143 out of 160 residues have been assigned. Additional residues were identified on the CACO spectrum but their amide protons could not be found, suggesting the amide protons were exchanging on a different timescale.

The extent of the main chain assignments is shown in figure 4.14, which highlights those residues for which an assignment could not be achieved. Clearly more of the 2-isobutylthiazole form of OBP3 was assigned than for unbound OBP3. Isolated residues that are shown as unassigned are generally due to missing peaks on 3D triple resonance experiments, where in most cases passes could be found from the neighbouring residues but the amide proton could not be found.

Unbound OBP3

MRGSHHHHHHTDPEEASFERNLDVDKLNWDWFSIVVASDKREKIEENGSMRVFVQHIDV
LENSLGFTFRKENGVCTEFSLVADKTAKDGEYFVEYDGENTFTILKTDYDNYVMFHLVN
VNNGETFQLMELYGRTKDLSSDIKEKFAKLCVAHGITRDNIIDLTKTDRCLQA

2-isobutylthiazole bound OBP3

MRGSHHHHHHTDPEEASFERNLDVDKLNWDWFSIVVASDKREKIEENGSMRVFVQHIDV
LENSLGFTFRKENGVCTEFSLVADKTAKDGEYFVEYDGENTFTILKTDYDNYVMFHLVN
VNNGETFQLMELYGRTKDLSSDIKEKFAKLCVAHGITRDNIIDLTKTDRCLQA

Figure 4.14 Residues for which NMR assignments are missing highlighted in green (unbound) and blue (2-isobutylthiazole bound).

For OBP3 bound to 2-isobutylthiazole a near complete set of $H\alpha$ chemical shifts were determined (the majority using the HACACO experiment). It was therefore possible to make an initial prediction of the secondary structure of OBP3. This could also be compared to the predictions made using CD in chapter 3 (section 3.2.4.2). In figure 4.15 the $H\alpha$ chemical shifts of the amino acid residues in a random coil (source BMRB) were subtracted from the assigned $H\alpha$ shifts. Values that are positive indicate β -strands whilst the negative values are indicative of helices. Clearly more helical character than would be expected from a lipocalin is seen (though less than predicted by CD). There are some dense collections of positive residues, indicating β -strands, between residues 22-26, 43-48, 54-60, 63-71 and 112-119. There is also a dense area of negative values between residues 133-147 which correlates to the possible location of a C-terminal α -helix.

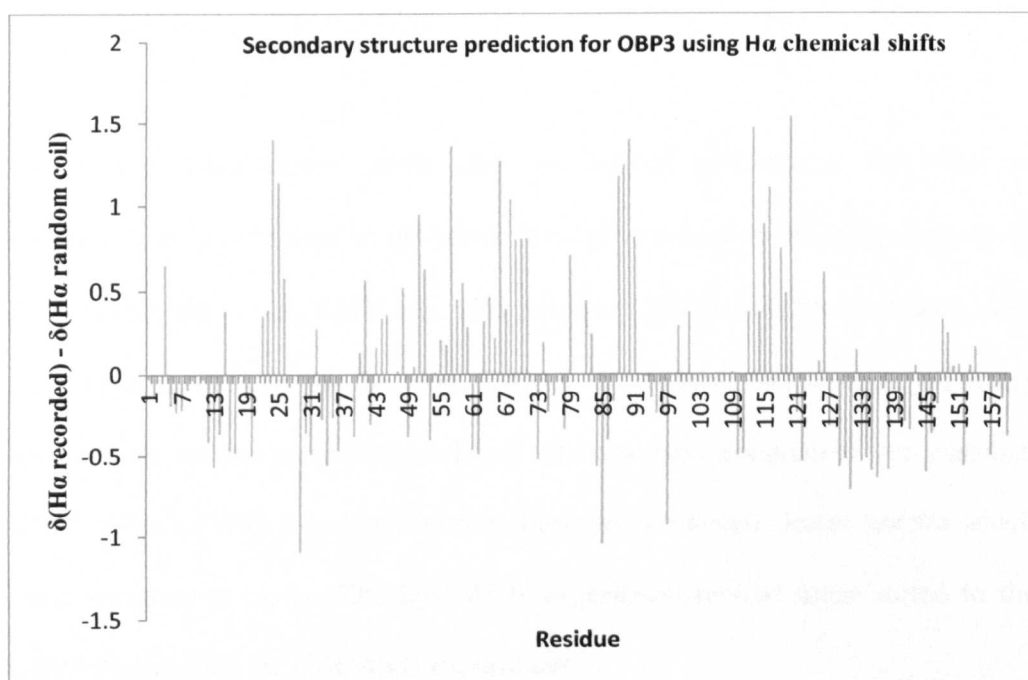


Figure 4.15 Predicted secondary structure of OBP3, using $H\alpha$ chemical shifts, positive values indicate β -strands, whilst negative values indicate helical character.

4.3 Conclusions

In order to conduct useful NMR experiments the spectrum of OBP3 needed to be assigned. No spectra of other vertebrate odour binding proteins have been assigned by NMR, however a number of major urinary proteins (MUPs) have been analysed (Francesco Abbate et al. 1999; Lücke et al. 1999; Phelan et al. 2010). OBP3 is considered to be “MUP-like” and in particular shares 65 % sequence identity with Darcin, which was recently assigned by NMR (Phelan et al., 2010). The $^1\text{H} / ^{15}\text{N}$ HSQC resembled that of OBP3, but didn’t overlay closely enough for an easy transfer of the assignment.

Darcin was able to be almost completely assigned, including β -strand G, which proved problematic when assigning OBP3, however the amino acid composition of this strand in Darcin is different; lacking the phenylalanine (Phe 103) and two valine residues (Val 106 and 108).

OBP3 is a promiscuous binder and this proved problematic for NMR as conducting an experiment in an odour free environment is virtually impossible. Even within the sealed NMR tube it is possible OBP3 could be interacting with the plastic in the cap on the tube. Taking all available precautions, relatively reproducible spectra were produced and near complete assignment was possible. OBP3 saturated with 2-isobutylthiazole, however, produced clearer spectra which made assignment easier. The HNCACB experiment proved more suited to the OBP3 system than the CBCANH experiment.

“Unlabelling” experiments proved useful and appeared to work well. The appearance of extra signals in the lysine unlabelled spectrum was surprising, and may have been due to some degraded protein. However, this wasn’t seen on any other spectra of OBP3. It was not possible to identify the additional signals. The disadvantage of “unlabelling” is that it requires a full round of protein production for each amino acid that is unlabelled. However it is considerably cheaper than selectively labelling particular amino acids.

The CACO and CCCO-TOCSY experiments proved very useful for identifying aliphatic amino acids and were particularly easy to use as they lacked artefacts from solvent encountered when using proton based experiments (due to the fact the carbon signals are directly observed and detected). A single carbon-13 only labelled sample was used as this relatively inexpensive to produce and nitrogen-15 labelling was not required.

The assignments achieved allows binding experiments on OBP3 to be performed, allowing study of OBP3 binding on a per residue basis, as will be described in chapter 5.

5. INITIAL BINDING CHARACTERISATION OF OBP3

5.1 Introduction

Previous work by Löbel et al. (2001, 2002) found a total of 25 odorous compounds that OBP3 was able to bind to. These studies were discussed briefly in section 1.4 in relation to the information they provided on the possible discriminatory properties of the rat OBP subtypes. Different chemical classes were assessed, including terpenes, pyrazines and thiazoles. It was suggested that each rat OBP showed a distinct ligand binding profile. In total, a set of 49 odorous compounds were screened for their ability to bind to each of the three rat OBP subtypes. This study will be examined more closely in this chapter in order to decide which odours to examine using isothermal titration calorimetry (ITC). This technique was chosen to look at the possible roots of rat OBP specificity and the nature of OBP binding in general.

In this chapter ITC provided information about the affinities and thermodynamics of ligand binding to OBP3. Ligand binding to OBP3 was then examined using NMR titrations; in order to map these interactions a structural model was required. This involved using high resolution structures that had good sequence identity to OBP3 on which to base a model of the structure of OBP3.

5.1.1 Isothermal titration calorimetry

In its most simple application isothermal titration calorimetry (ITC) can be used to determine the affinities of binding interactions, which, importantly, can be measured directly. However, it is a valuable technique as thermodynamic details of the reaction are also directly determined, without the need for van't Hoff analysis, avoiding the temperature dependence of the free energy change or affinity.

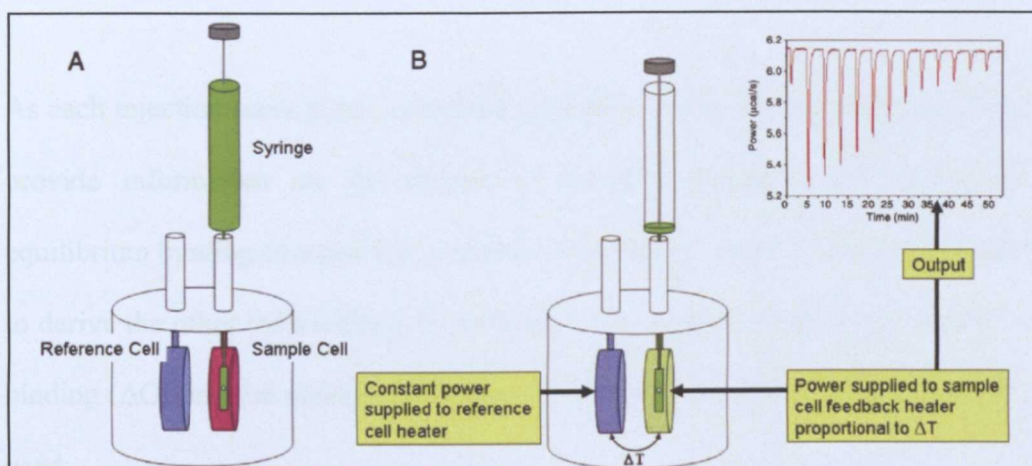


Figure 5.1 Schematic diagram of an isothermal titration calorimeter. “A” shows the basic instrument layout with the reference and sample cells contained within temperature controlled jacket. The syringe shown in green (with a paddle at the bottom for stirring), contains the ligand solution, which is being injected into the macromolecule solution (pink) in the sample cell. “B” shows that as the reaction takes place (the solution in the sample cell is shown as pale green) a heat change takes place and the power required to adjust the sample cell back to the same temperature as the reference cell is plotted as a series of injection. (Ladbury 2004)

The schematic diagram in figure 5.1 shows an ITC instrument and the stages involved in data acquisition. The instrument contains two cells within an external jacket; the reference cell and the sample cell, in a similar manner to that described for DSC in section 3.2.5.2 In ITC the reference cell remains at a constant temperature. The solution in the sample cell is titrated with aliquots (each referred to as an injection) of the reactant which cause a heat change. The recorded parameter is the power required to return the sample cell to the same temperature as the reference cell. The apparent enthalpy of binding (ΔH) can be calculated by integrating the power required with respect to time.

As each injection takes place, saturation gradually occurs and the integrated heats provide information on the degree of complex formation and hence the equilibrium binding constant K_a (a measure of affinity) can be calculated. In order to derive the other thermodynamic parameters; the observed Gibbs free energy of binding (ΔG) and the observed binding entropy (ΔS), the following equations are used:

$$\Delta G^o = RT \ln K_a \quad (\text{Equation 5.1})$$

Where R is the gas constant = $8.31 \text{ J K}^{-1} \text{ mol}^{-1}$ and T is the absolute temperature in K. The free energy of binding is then related to the enthalpy and entropy terms by:

$$\Delta G^o = \Delta H^o - \Delta S^o \quad (\text{Equation 5.2})$$

If experiments are measured at a number of temperatures the heat capacity (ΔC_p) of the binding reaction can be determined using the following equation

$$\Delta Cp = \frac{\delta(\Delta H^o)}{\delta T} \quad (\text{Equation 5.3})$$

Where δT is the temperature range over which the enthalpy change is measured. Heat capacity can be used to relate thermodynamic data to structural details. It is important to note that the values determined by ITC should be treated as “apparent”. This is because the heats determined not only involve the heat of binding but also events such as protonation, solvation and conformation transitions. Blank titrations (i.e. ligand titrated into buffer alone) are conducted to reduce the effect of these phenomena.

ITC has many advantages, particularly that it measures the heat with extremely high sensitivity, as low as several hundred nanojoules can be recorded (Falconer et al., 2010) and it is a very accurate technique. Additionally ITC does not require probes to be attached to the macromolecule or ligand to monitor the reaction, or for the macromolecule to be immobilised, as is required by other techniques that monitor binding interactions. The experiments are also performed in biologically relevant, aqueous buffers.

There are some drawbacks to the technique; ITC requires large amounts of sample when compared to other techniques such as surface plasmon resonance and fluorescence assays. Another limiting factor of ITC is the range of reaction affinities that can be measured, which is between $\sim 1 \text{ nM} < K_d < \sim \text{several hundred micromolar}$ (Falconer et al., 2010). Very weak interactions are difficult to measure because the reagent concentrations required to reach saturation are

prohibitive in terms of cost and sample solubility. High affinity interactions can tend to appear as “step” functions, with no or very few data points in the transition of the binding curve. Using a low concentration of sample can alleviate this but causes low signal to noise and can pass below the instrument’s limit of detection.

The experiments in this chapter were all carried out using 30 μM protein and 300 μM ligand. This enabled rapid screening and a broad gauge of which ligands OBP3 bound tightly or weakly to. The range of dissociation constants (K_d) that can be measured accurately is limited to between 3 μM (highest affinity) and 30 μM (lowest affinity). Beyond these limits the ITC curve fitting becomes less than ideal and the dependent parameters of ΔH , ΔS and ΔG become more inaccurate and should be treated with caution. Calculation of the parameter K (binding constant) and thus the other values are dependent on an accurate stoichiometry and therefore if the protein or ligand concentration were not as expected the values obtained would become incorrect. The protein concentration was measured using UV spectroscopy and an accurately determined protein extinction coefficient (section 2.5.3). The ligand concentration was more difficult to measure accurately, density measurements were used and care was taken to ensure the ligand was completely solubilised (section 2.6.2).

ITC is a very powerful technique and has been used to look at interactions including; protein-protein, protein-small molecule (such as drugs), protein-metal, protein-nucleic acid, protein-lipid, nucleic acid-small molecule and enzyme activity (Falconer et al., 2010). To be particularly useful in drug design the

relationship between calorimetric data and molecular structure is a key consideration. Spotting trends and relationships is a difficult task and to this end attempts have been made to create calorimetric databases (Liu et al. 2007, 2008) for the correlation of data.

5.1.2 Selection of odorants

In order to decide which odours to examine by ITC, the data for odours previously determined to bind to OBP3, along with the 25 odours found not to bind to OBP3 were examined closely. It should be noted that Löbel et al. (2002) used fluorescence competition assays. A fluorescent probe, (in the case of OBP3, 1-amino-anthracene (1-AMA)) was bound to the OBP and odour compound added whilst the change in relative fluorescence was monitored as the probe was displaced. This means that only odour compounds with higher affinity for the OBP than the probe could displace the latter. Table 5.1 is reproduced from the study and shows OBP3 to have the lowest affinity fluorescent probe (1.22 μ M dissociation constant).

	Rat OBP1	Rat OBP2	Rat OBP3
Fluorescent probe	1-AMA	1,8-ANS	1-AMA
Chromophore net charge	positive	negative	positive
Emission maximum (nm)	502	462	480
Dissociation constant (μM)	0.59 ± 0.03	0.62 ± 0.05	1.22 ± 0.08

Table 5.1 Dissociation constants of Rat OBPs and the fluorescent probes used reproduced from Löbel et al. (2002)

This may explain why it appears that OBP3 is the most promiscuous binder, as even weaker ligands could displace the 1-AMA probe. Tables 5.2 and 5.3 show that OBP3 is able to bind to approximately 20 % more odours than OBP1. OBP2 showed binding to only 5 out of the 49 odours (20% fewer than OBP1). The affinity values given in the tables are IC_{50} and K_i which indicate that these are competition experiments. The IC_{50} amount of ligand required to displace half of the fluorescent probes, and is the ligand concentration measured at 50 % of the maximum fluorescence emission curve. The term is commonly used in antagonist drug potency studies, and is a measure of the half maximal inhibitory concentration of the antagonist (in this case the antagonist is the added odour). K_i is the absolute inhibitory contribution of the antagonist (or competitive agonist) and, unlike IC_{50} , is a direct report of affinity (a dissociation constant), with a smaller number indicating higher affinity. K_i is calculated from IC_{50} by dividing it by 1 plus the substrate concentration (in this case the fluorescent probe concentration) over K_m . In this case K_m was calculated using the fluorescent probe dissociation constants.

Odour	Molecular weight	IC ₅₀ (μM)	Ki (μM)
2,6-Isopropylphenol	178.27	0.58	0.16
2-Isobutylthiazole	141.23	2.03	0.58
2,4,5-Methylthiazole	127.13	3.05	0.87
2,5-Dimethylpyrazine	108.14	4.62	1.32
2,3,5 -Trimethylpyrazine	122.17	5.42	1.55
2,3-Diethylpyrazine	136.2	5.85	1.67
Benzothiazole	135.19	5.84	1.67
2,3-Dimethylpyrazine	108.14	6.3	1.8
2,6-Dimethylpyrazine	108.14	6.3	1.8
4,5-Methylthiazole	113.18	6.62	1.88
Keton-mosculus (Musk ketone)	294.31	6.62	1.88
2,3,5,6-Tetramethylpyrazine	136.19	6.8	1.94
2-Ethylpyrazine	108.14	8.75	2.15
(-)-Menthone	154.25	7.7	2.21
D-(-)-Limonene	136.24	8.43	2.41
2-Isobutyl-3-methoxypyrazine	166.22	10.1	2.88
(+)-trans-Limonene-1,2 - epoxide	152.23	10.15	2.9
(+)-cis Limonene- 1,2-epoxide	152.23	10.22	2.92
5-Methylthiazole	99.15	13.45	3.85
(-)-Menthole	156.27	14.3	4.12
2-Acetylthiazole	127.164	16.52	4.72
2-Methylpyrazine	94.11	24.42	6.97
Myristyl amine	*	24.2	7.56
4-Methyl-5-ethanol-thiazole	143.21	27.95	7.98
2-Methyl-3-methoxypyrazine	124.14	28.2	8.05

Table 5.2 Odours found to bind to OBP3 (Löbel et al. 2001, 2002) Listed in order of affinity.

Coloured by chemical class (Red = thiazole, Green = pyrazine, blue = monocyclic terpenes, orange = phenols. *Indicates that the exact odour used could not be identified.

Odour	Molecular Weight	Bind to OBP1	Bind to OBP2
(-)-Borneol	154.25	1.25	
(-)-Camphor	152.23	1.35	
(+)-Camphor	152.23	1.29	
Phenol	94.11		
4-Ethylphenol	122.16	3.69	
4-Propylphenol	136.19	0.54	
4-Isopropylphenol	136.19		
3-Isopropylphenol	136.19	0.52	
2-Isopropylphenol	136.19		
Benzaldehyde	106.12		
1-Decanol	158.28		
1-Decanal	156.2		1.24
Capric acid	172.26		
Myristyl alcohol	214.39		
Myristine aldehyde	212.37		0.41
Myristic acid	228.37		0.29
3,7-Dimethyl-2,6-octadiene nitrile	149.23		2.42
p-tert-Butyl- α -methyl dihydrocinnamic aldehyde	204.3		1.77
Pyrazine	80.09		
2-Acetylpyrazine	122.12		
2-Methoxypyrazine	110.11		
Thiazole	85.13		
2-Ethoxythiazole	129.18		
4-Methyl-5-acetic acid-ethylester-thiazole	185.24		

Table 5.3 Odours found not to show an interaction with OBP3 (Löbel et al., 2002). Those shaded grey do not bind to any OBP. Those that bind to OBP1 or OBP2 have a K_i value listed in the table.

Generally it was found that OBP3 was able to bind well to heterocyclic compounds, the thiazoles and pyrazines. OBP1 was only shown to be able to bind to 2-isobutyl-3-methoxypyrazine and 2-isobutylthiazole from these groups. These two odours were therefore selected for closer study to allow comparison between OBP1 and OBP3 and to establish the thermodynamic basis of binding of these compounds to OBP3. 4,5-dimethylthiazole along with 2-isobutylthiazole had previously been tested by ITC (Löbel et al. 2001) and therefore this thiazole was also included for comparison (Figure 5.2).

	<p>Name: 2-isobutyl-3-methoxypyrazine</p> <p>Description: Pyrazine</p> <p>Molecular weight : 166.22</p> <p>Log P : 2.62 (estimated)</p> <p>Odour: Bell pepper</p>
	<p>Name: 2-isobutylthiazole</p> <p>Description: Thiazole</p> <p>Molecular weight : 141.2</p> <p>Log P : 2.51 (estimated)</p> <p>Odour: Tomato leaf</p>
	<p>Name: 4,5-dimethylthiazole</p> <p>Description: Thiazole</p> <p>Molecular weight : 113.2</p> <p>Log P : 2.09 (estimated)</p> <p>Odour: Fishy</p>

Figure 5.2 Thiazoles and pyrazines tested by ITC. The partition coefficient of the odour is described by: $\text{Log } P = \log \left(\frac{[\text{solute}]_{\text{octanol}}}{[\text{solute}]_{\text{un-ionised water}}} \right)$ the larger Log P is the more hydrophobic the

molecule. The calculated (estimated) value may be skewed by presence of polar groups on the molecule.

Interestingly 2,6-diisopropylphenol was shown to have the highest affinity interaction with OBP3, despite being the only phenol derivative that OBP3 bound to. It was also the only phenol derivative that OBP1 didn't show binding to. Propofol is the drug name of 2,6-diisopropylphenol, and it is used as a general anaesthetic and sedative. It isn't noted for its odorous properties and in fact isn't very volatile. It is noted to be readily bound by proteins such as albumin (Favetta et al., 2002). Due to a lack of availability and the potential harmful nature of this compound it wasn't tested by ITC. Löbel et al. (2002) suggest 2,6-diisopropylphenol is able to bind to OBP3, whilst the other phenol derivatives are unable to, because the isopropyl groups at positions 2 and 6 provide a steric block to the hydroxyl group at position 1. Additionally it was noted that the aliphatic alcohols studied (such as decanol) did not demonstrate binding to OBP3. Therefore three compounds with hydroxyl groups in varying configurations were chosen for further testing. These are shown in the figure 5.3.


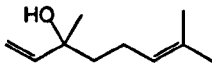
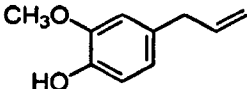
	<p>Name: Octan-1-ol</p> <p>Description: Straight chain alcohol</p> <p>Molecular weight : 130.2</p> <p>Log P : 2.88 (estimated)</p> <p>Odour: Orange</p>
	<p>Name: (R)-(-) Linalool</p> <p>Description: Branched chain, terpene alcohol</p> <p>Molecular weight: 154.2</p> <p>Log P: 3.28 (estimated)</p> <p>Odour: Lavender</p>
	<p>Name: Eugenol</p> <p>Description: allyl chain-substituted guaiacol</p> <p>Molecular weight: 164.2</p> <p>Log P: 2.27</p> <p>Odour: Clove</p>

Figure 5.3 Hydroxyl containing odours tested by ITC.

Octan-1-ol was selected as it is a straight chain alcohol like decanol, but it is smaller and therefore may be more readily accommodated into the binding site of OBP3. It was also contrasted with octan-1-al, the aldehyde equivalent of octan-1-ol to see if there was a positive or negative effect afforded by the hydroxyl group. Linalool has a slightly more sterically blocked tertiary alcohol. It is a shorter molecule than octanol, though overall it is a larger, albeit relatively simple, molecule. It has a stereogenic centre at C₃ and therefore can exist as either an R or S enantiomer. Both forms are present in nature, the R form possesses a woody lavender aroma whilst the S form is more sweet and floral (Ulland et al., 2006). The R form was used for these studies. Eugenol, the final hydroxyl containing

compound tested, is slightly larger again and contains a phenolic hydroxyl. The hydroxyl group is somewhat occluded by a methoxy group attached to the neighbouring ring carbon.

Finally ethyl butyrate and γ -decalactone were included in the ITC screen. Ethyl butyrate is a straight chain ester with a lone oxygen adjacent to the double bonded oxygen, whilst lactones, such as γ -decalactone are cyclic esters which have a double bonded oxygen attached to a ring which has an adjacent lone oxygen. These compounds have been found to bind to OBP1 in previous studies in this laboratory and elsewhere (Nespoulous et al. 2004), therefore comparisons could possibly be made.

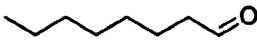
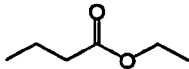
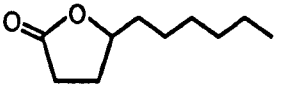
	<p>Name: Octanal</p> <p>Description: Straight chain aldehyde</p> <p>Molecular weight : 128.2</p> <p>Log P : 2.95(estimated)</p> <p>Odour: Citrus peel</p>
	<p>Name: Ethyl butyrate</p> <p>Description: Straight chain ester</p> <p>Molecular weight : 116.2</p> <p>Log P : 1.85 (estimated)</p> <p>Odour: Sweet fruity</p>
	<p>Name: γ-decalactone</p> <p>Description: Cyclic ester</p> <p>Molecular weight : 170.2</p> <p>Log P : 2.72</p> <p>Odour: Fruit- peachy</p>

Figure 5.4 Odours containing a double bonded oxygen atom tested by ITC.

5.2 Results and Discussion

5.2.1 Thiazole and pyrazine ITC binding experiments

As discussed in section 5.1.1, a ligand injection causes a heat change (shown by a single spike in the raw ITC traces in the top panels of figure 5.5). A series of injections (20 injections of 14.03 μ l in these experiments) eventually leads to saturation of all the protein binding sites by the ligand and no further heat change takes place. In the case of the following experiments the heat changes (ΔH°) are all negative, showing the reactions are exothermic.

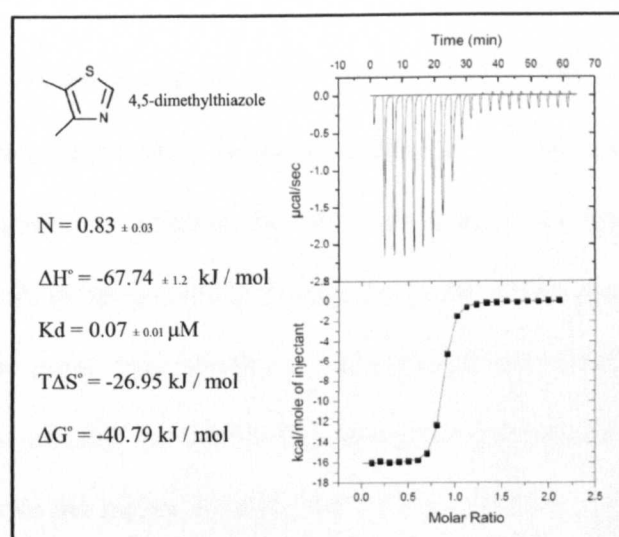
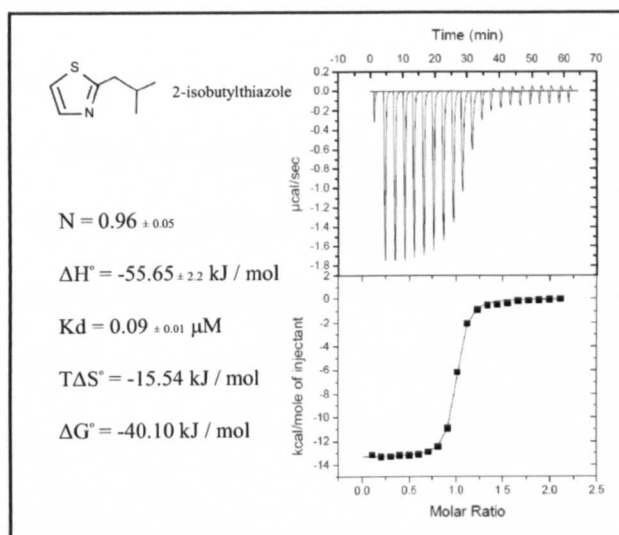
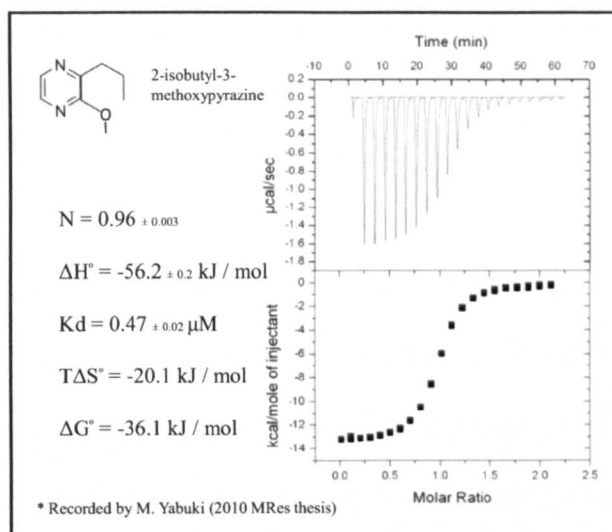


Figure 5.5 ITC results of the thiazoles and pyrazines, raw injection data are in the top right, the fitted data is the bottom right, whilst the thermodynamic parameters are in the bottom left. The

calorimeter records the data in calories however for the purpose of analysis the values were converted into kJ. Errors quoted are the standard error of the mean of duplicate experiments.

In figure 5.5 the bottom panel shows the plot of the integrated heats and a best fit line through the points. The parameter n describes the number of binding sites, giving the stoichiometry of the interaction. The expected n value would be 1.0; however the values range from 0.83 to 0.96 which is indicative of there being a degree of inactive protein (either being incorrectly folded, non-functional or already having ligand bound after purification). The n values found by Löbel et al. (2001) were lower for 4-5-dimethylthiazole (DMT) and 2-isobutylthiazole (IBT) (0.76 and 0.78). This may be due to using a different method of endogenous ligand removal (section 3.2.2.4) that possibly wasn't as effective or it could be that they used lower purity ligands, affecting the stoichiometries calculated. Alternatively it could be the case that there was more inactive protein.

Similar to the results found by Löbel et al. (2001, 2002) the heterocyclic compounds tested were tightly bound to OBP3. In fact the thiazoles bound so tightly (with nanomolar affinity) that the transition of the binding curve was somewhat difficult to define (being composed of only 3 data points), although the fits appear to be good. Interestingly 4-5-dimethylthiazole (DMT) was found to have the highest affinity for OBP3, this is contrary to Löbel et al. (2001) who found IBT to have the higher affinity (0.47 μM for IBT vs. 1.89 μM for DMT). There is also a considerable difference in the magnitude of the binding affinity. (This study found values 5 times higher for IBT and 27 times higher for DMT).

The experiments in this study were conducted at 25 °C in Phosphate Buffered Saline (PBS) at pH 7.4 (described in section 2.6.1.), whereas in the study by Löbel et al. (2001) the ITC was performed at 30 °C in 100 mM potassium phosphate at pH 7.5. Experimental temperature and buffer conditions can affect the binding parameters and could be the source of the difference. However the macromolecules are far in excess and therefore the protonation effects of the buffers are very small, therefore it is most likely the differences is due to differing degrees of active protein in the preparations.

Binding of IBT to OBP3 by ITC was run with 40 ligand injections of 7 µl in order to further define the binding curve transition for this ligand and get a more accurate binding affinity. Figure 5.8 shows the ITC result in the same format as figures 5.5 to 5.7. The curve is very similar to the initial 20 injection experiment. However the binding affinity measured is stronger (0.03 µM rather than 0.09 µM), the enthalpy value appears to be more negative, however the entropy value is also more negative, resulting in a similar free energy overall and the same enthalpy driven mode of ligand binding.

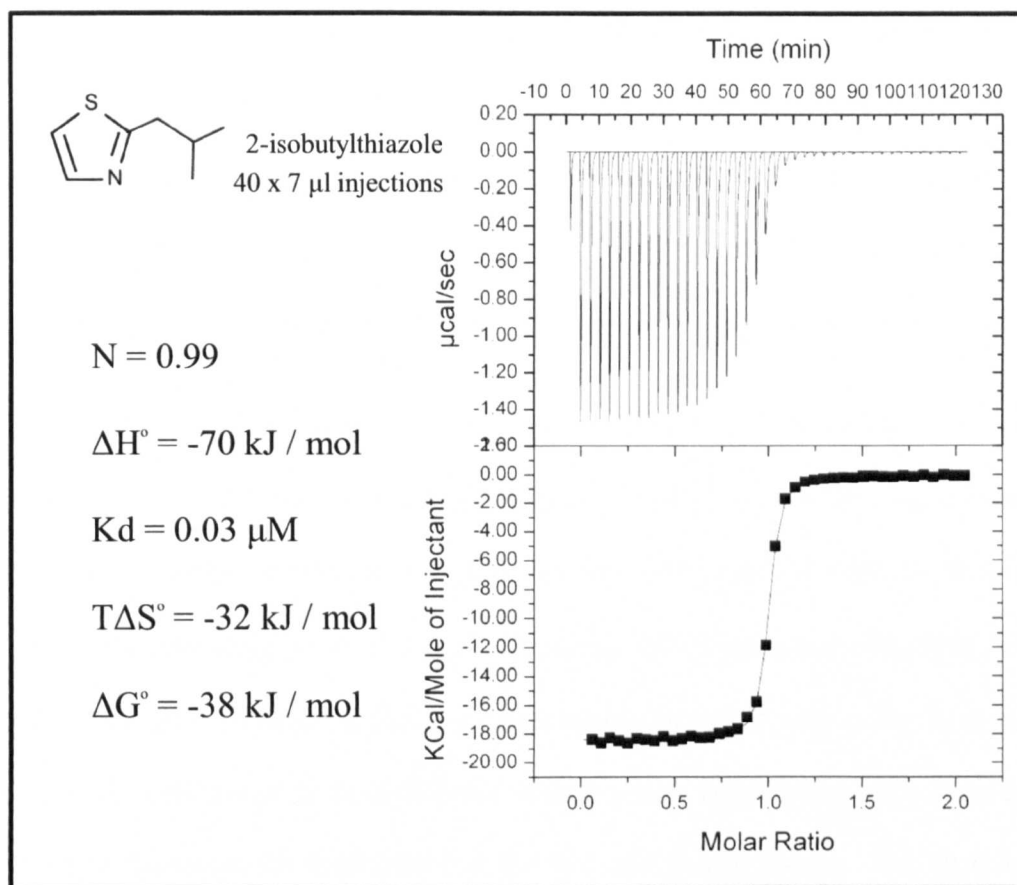


Figure 5.8 2-isobutylthiazole binding to OBP3 by ITC, with 40 injections (7 μ l each)

2-isobutyl-3-methoxypyrazine (IBMP) bound more weakly to OBP3 than the thiazoles tested. Löbel et al. (2002) also found this to be the case. However, the binding constant found by ITC in this study is again smaller (6 times lower) showing a higher affinity than found previously.

In terms of the thermodynamic parameters, the reactions are highly favourable with negative free energies of binding (with ΔG° values from -36.1 kJ/mol to -40.8 kJ/mol. Generally ΔG° can be considered to represent the following

contributions; intermolecular contacts (non-covalent bonds), conformational changes, hydrophobic transfer (of the ligand from the solvent into the binding site), ionization effects and changes in rotational and translational motions of all reactions components (Thomson & Ladbury 2004). It provides a concentration independent way of comparing affinities, in this case across a group of ligands.

From equation 5.2 in section 5.1.1 we can see that ΔG° can be derived from the enthalpy (ΔH°) and entropy terms (ΔS°) terms. The ligand with the largest (most negative) enthalpic contribution to binding was DMT (-67.74 kJ/mol). Despite having the least negative free energy of binding, IBMP has a more negative ΔH° value than IBT. Interpreting ΔH° , the heat energy involved with going from the free to the bound state is more difficult without a high resolution protein structure to relate the molecular interactions in the free and bound state to. The breaking and making of non-covalent bonds when forming the complex results in the net heat energy recorded. These bonds include those involved in ligand binding, and solvent rearrangement upon binding, as well as those involved in conformational changes of the ligand and macromolecule (Henriques & Ladbury 2001; Thomson & Ladbury 2004; Olsson et al. 2008). It is interesting to note that despite being derived from the same parent molecule (the thiazole ring) there is a 12 kJ/mol difference in ΔH° resulting from the positions and presence of methyl groups at positions 4 and 5 in DMT rather than an isobutyl at position 2 found in IBT.

The entropy term (ΔS°) for IBT is the least negative and therefore gives the least unfavourable contribution to the free energy of binding. In terms of ligand binding there are two sets of processes to consider as ΔS° reflects whether the system

becomes more ordered (giving a negative ΔS° value) or disordered (giving a positive ΔS° value) when the ligand and macromolecule come together. Order is imposed through translational entropy and a restriction in the atomic degrees of freedom from bonding in the binding interface (configurational entropy). The release of water, such as when a hydrophobic surface is buried, results in an increase in the disorder of the system. It may therefore be suggested that the binding of these ligands increases the order of the system through the restriction of the degrees of freedom of the protein and ligand. When binding takes place not enough ordered water is released upon binding to cause an increase in disorder and a positive entropy term.

5.2.2 ITC binding of ligands containing a hydroxyl group

The ITC plots and binding curves of the odours tested which contained hydroxyl groups; octan-1-ol, linalool and eugenol are shown in figure 5.6 using the same format as in figure 5.5. The n values were close to 1 for the three ligands, suggesting a 1:1 binding stoichiometry like that seen for the previous ligands. These three ligands were more weakly bound to OBP3 than the heterocyclic compounds (between 9 and 140 times weaker), having micromolar rather than nanomolar affinities. Despite the variation in both the size and shape of the molecules their binding profiles are quite similar to each other with little variation in binding affinity ($K_d = 2.41$ to $9.8 \mu\text{M}$) and free energy ($\Delta G = -28.57 \mu\text{M}$ to -30.56 kJ/mol). As suggested by Löbel et al. (2002), who found that decanol didn't show any binding to OBP3, the hydroxyl group is not well tolerated by OBP3. This may be due to a lack of side chains in the protein or water for the OH group to hydrogen bond with.

The entropy values for all three odours are positive ($\Delta S^\circ = 4.08$ to 10.32) showing an increase in entropy upon binding. Octan-1-ol has the least positive entropy change, however it has the most negative enthalpy change, this may suggest a number of binding contacts are being made or that this straight chain molecule is restricted in the number of ways it can occupy the binding site of OBP3. The effect of the length of the straight chain in a molecule on binding thermodynamics is explored further in chapter 6.

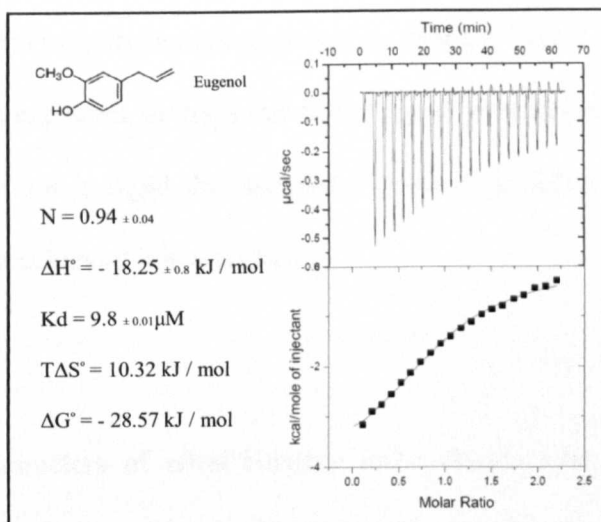
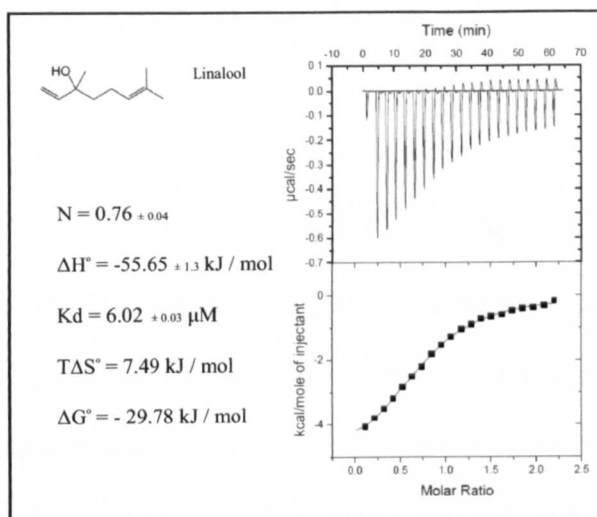
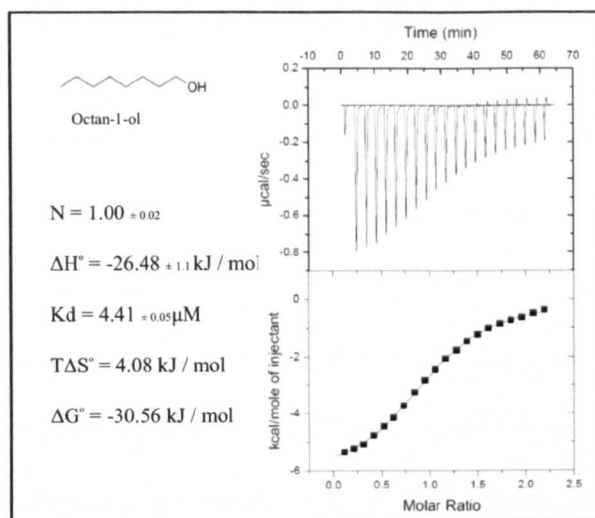


Figure 5.6 ITC of odours containing an hydroxyl group. Errors quoted are the standard error of the mean of duplicate experiments.

5.2.3 ITC binding of molecules containing a double bonded oxygen atom

It is somewhat surprising that whilst octan-1-ol is able to bind to OBP3 simply replacing the hydroxyl group with an aldehyde group reduces the binding drastically. Whilst some interaction is seen, the binding curve of octanal binding to OBP3 is not possible to fit (figure 5.7). This result matches the finding of Löbel et al (2002) that decanal does not bind to OBP3. It also suggests that the lack of binding is not purely due to the length of the aliphatic chain, although it might be a further hindrance that abolishes binding altogether.

Ethyl butyrate and γ -decalactone (figure 5.7) were found to have micromolar binding affinity to OBP3 (1.98 and 2.12 μM respectively). γ -decalatone had an n value of 0.96 which is very close to the expected 1:1 binding, however ethyl butyrate caused saturation to occur very quickly and it is unclear why this should be the case. A higher concentration (600 μM of ligand rather than 300 μM) was used for these experiments; however, it may be that the ligand concentration was even higher than expected, perhaps due to the ligand preparation being purer than stated. As saturation is rapid the sigmoidal curve is ill defined and although a good fit has been achieved it is not ideal.

The binding parameters of ethyl butyrate and γ -decalactone binding are very similar, with near identical free energies of binding ($\Delta G^\circ_b = \sim 32 \text{ kJ/mol}$) and the change in entropy is also close ($\Delta S^\circ_b = -15.2 \text{ kJ/mol}$ and -12.4 kJ/mol respectively). There is a slight difference in the enthalpies of binding, but they

were both highly negative ($\Delta H^\circ_b = -47.70$ and -44.77 kJ/mol respectively). These parameters suggest that both ligands bind in a manner that is enthalpy rather than entropy driven. This is interesting as γ -decalactone has a substantial aliphatic chain to be accommodated within the binding site, which in the cases of octan-1-ol and octanal appeared to adversely affect binding. For octan-1-ol this appeared to be due to a less negative enthalpic contribution than other odorants despite a positive entropy value. In chapter 6 the binding of a series of γ -lactones is explored in an attempt to further clarify the thermodynamics of binding.

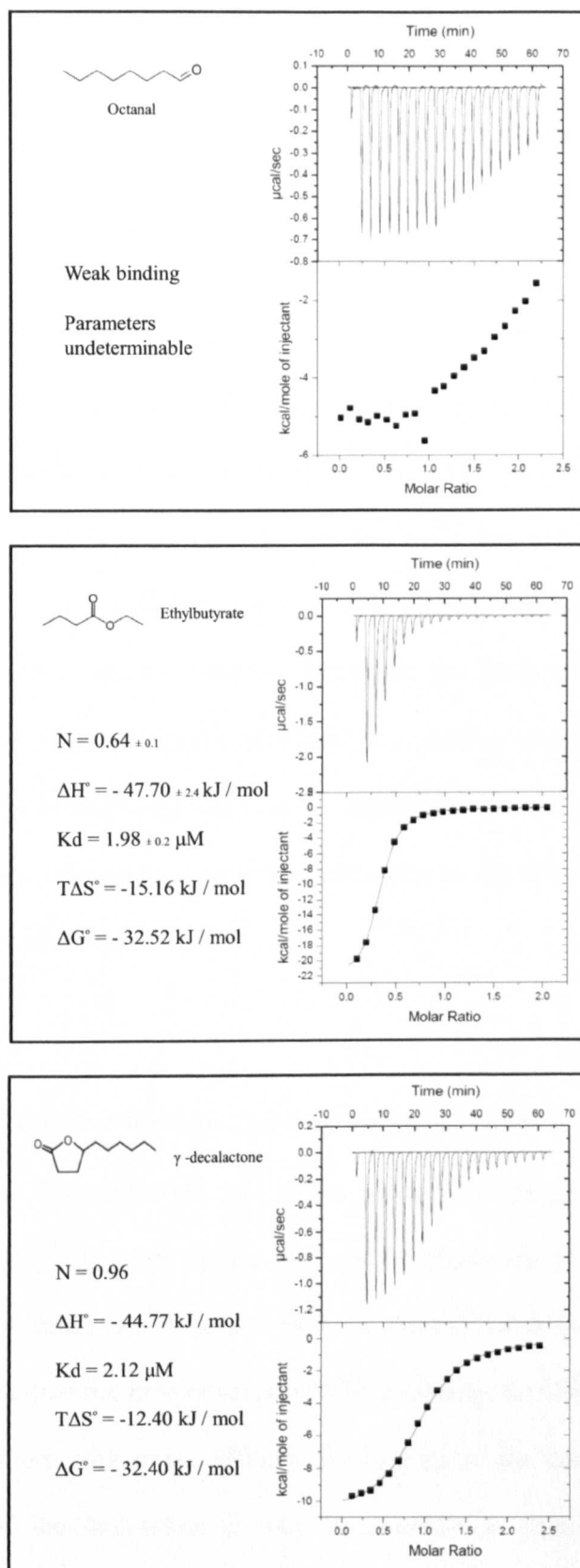


Figure 5.7 ITC of odours containing a double bonded oxygen. Errors quoted are the standard error of the mean of duplicate experiments.

5.2.4 Ligand binding by ITC

Looking at all nine odours, patterns are difficult spot. Comparing the affinity values with Log P or the molecular weights of the odours does not reveal an obvious correlation across all the odours. However, for the heterocyclic compounds it appears that the bigger the molecular weight and Log P value, the weaker the binding affinity is. For the hydroxyl group the higher molecular weight odours also have the lowest affinities. The Log P values don't follow this pattern. Ethylbutyrate is smaller and less hydrophobic than γ -decalactone and has higher affinity. Octanal falls between these two odours, however it has very low binding. These observations therefore highlight the importance of functional groups in determining the binding affinities. It is curious that, despite lipocalins being reported as hydrophobic binders, the more hydrophobic odours seemed to bind more weakly, this may however be a reflection of the effect of molecule size rather than hydrophobicity.

Overall it was found that the thiazoles were the highest affinity ligands for OBP3 of those tested. The affinities are in the nanomolar range rather than the micromolar range. This raises the question; is the physiological role of OBP3 to specifically bind small heterocycles? When a mixture of odorants is present it could be the case that the heterocycles with high affinity for OBP3 would always out-compete odours with lower affinity. By looking at the contributions to the high affinity of the interaction it may be possible to establish if OBP3 is specifically adapted for these ligands.

Although 4-5-dimethylthiazole (DMT) was found to be the highest affinity ligand the bound assignment of OBP3 had already been undertaken on 2-isobutylthiazole (IBT) and so it was sensible to carry out NMR titrations using this ligand. Additionally IBT was available in higher purity than DMT(>99% rather than >97%), which meant it was less likely to have artefacts from contaminants that may interfere with NMR than DMT.

5.2.5 Constructing a homology model of OBP3

A number of odour binding proteins, major urinary proteins and other lipocalins have high resolution structural entries in the Protein Data Bank (PDB) suggesting that it be possible to construct a good quality homology model of OBP3. In order to find a template on which to model OBP3, a BLAST search (Altschul et al., 1990) was conducted on the primary amino acid sequence of OBP3 against all proteins in the PDB. The closest sequence identity found was to α_{2u} -Globulin (A2U, PDB code: 2A2U), which is unsurprising as OBP3 was initially identified as having a very similar nucleotide sequence. OBP3, however, was noted to be expressed into the nasal cavity rather than into the urine as A2U is (Chaudhuri et al. 1999; Löbel et al. 2001). Alignment of the amino acid sequences shows them to have 87 % sequence identity, when the pre-sequences are excluded (figure 5.9).

```

OBP3      EEASFERGNLDVDKLNQDWFSIVVASDKREKIEENGSMRVFVQHIDVLENSLGFTFRIKE 60
A2U       EEASSTRGNLDVAKLNQDWFSIVVASNKREKIEENGSMRVFMQHIDVLENSLGFKFRIKE 60
          ****          *****          *****          *****          *****

OBP3      NGVCTEFSLVADKTAKDGEYFVEYDGENFTILKTDYDNYVMFHLVNVNNGETFQLMELY 120
A2U       NGECELYLVAYKTFEDGEYFVEYDGGNTFTILKTDYDRYVMFHLINFKNGETFQLMVLY 120
          ** * *: ** * *: ***** ***** ***** *: *: ***** **

OBP3      GRTKDLSSDIKEKFAKLCVAHGITRDNIIDLTKTDRCLQA-- 160
A2U       GRTKDLSSDIKEKFAKLCEAHGITRDNIIDLTKTDRCLQARG 162
          *****          *****

```

Figure 5.9 Alignment of the primary sequence of OBP3 and A2U. The proteins align well and have 87 % sequence identity.

The structure of A2U was solved as a tetramer (Bocskai et al., 1992; Chaudhuri et al., 1999), thus far no evidence of OBP3 being a tetramer has been found.

The crystal structure of A2U shows 4 monomers with 222 symmetry. There is no covalent bonding between the subunits. Each monomer is shown to bind to a single ligand molecule. It is unclear if this is a true tetramer, or a crystallographic artefact as most lipocalins are not multimeric. It should also be noted that other OBP x-ray crystal structures have found closely packed protein monomers in the crystal unit but no actual subunit contacts (White et al., 2009).

The homology model of OBP3 as a tetramer modelled on A2U using Swiss Model homology modelling server (Guex & Peitsch 1997; Schwede et al. 2003; Arnold et al. 2006; Kiefer et al. 2009) is shown in figure 5.10. As expected the model matches closely the template model A2U. The four monomers are rotated 180° to each other. The monomers possess the typical lipocalin fold, an antiparallel β -barrel with an α -helix at the C-terminal.

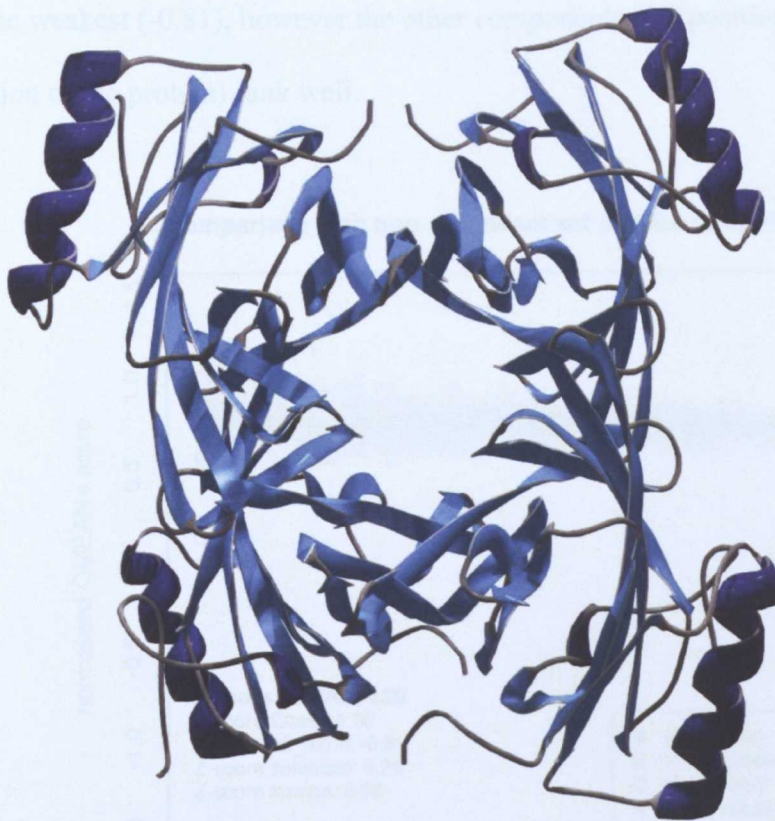


Figure 5.10 Homology model of OBP3 using A2U as a template, reveals a tetrameric molecule.

The quality of the model was assessed using the standard Swiss Model estimates of quality; for global quality the QMEAN4 scores (Benkert et al., 2011) and for local quality QMEAN and ANOLEA (Melo et al., 1997). The overall QMEAN4 score was 0.8. This is a measure of overall quality used to predict the reliability of the model and compare different models of the same target. The scale ranges from 0 to 1 (unreliable to reliable), suggesting, on a whole, the quality of this model was good. Figure 5.11 show the QMEAN Z-score, which overall was 0.29. This ranks the model quality against a reference set of high resolution structures, on average high resolution x-ray crystal structures have a value of 0. Also listed are the component parts of the Z score. The “all atom interactions” part of this score

is the weakest (-0.81), however the other components (C β positions, solvation and torsion of the protein) rank well.

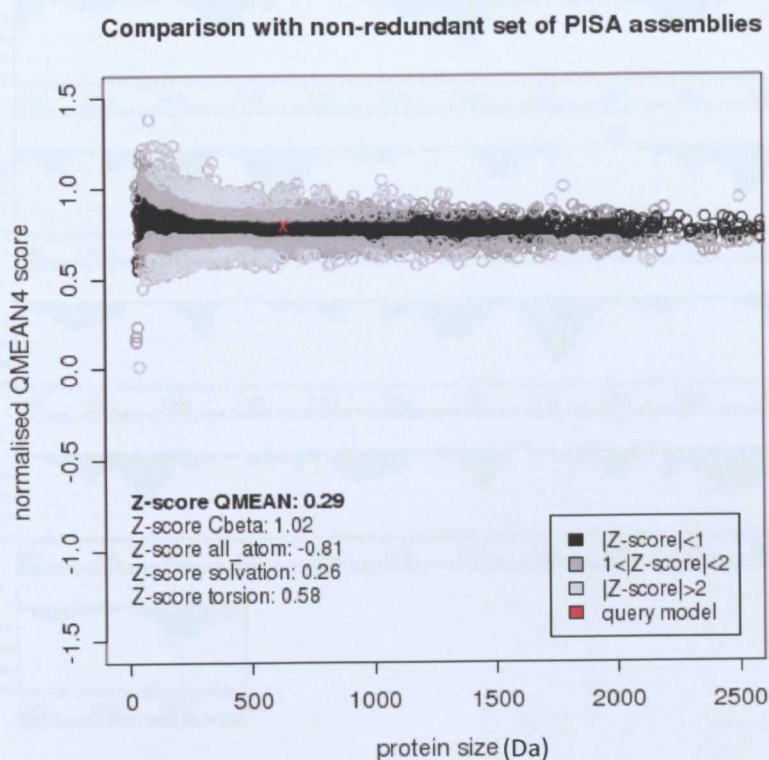


Figure 5.11 The QMEAN4 score of OBP3 modelled on A2U (red cross) plotted with a reference set of crystal structures.

ANOLEA was used to assess the acceptability of the local environments of the amino acids. This is shown in figure 5.12. Generally most amino acids were found to have favourable energies (giving negative values and appearing as green). The N-terminus was the major unfavourable region, along with 3 residues (74-76) at the end of β -strand D (the fourth strand from the N-terminus). In the template molecule there is a proline at position 75, but OBP3 has an alanine, this may account for the difficulty in modelling this position.

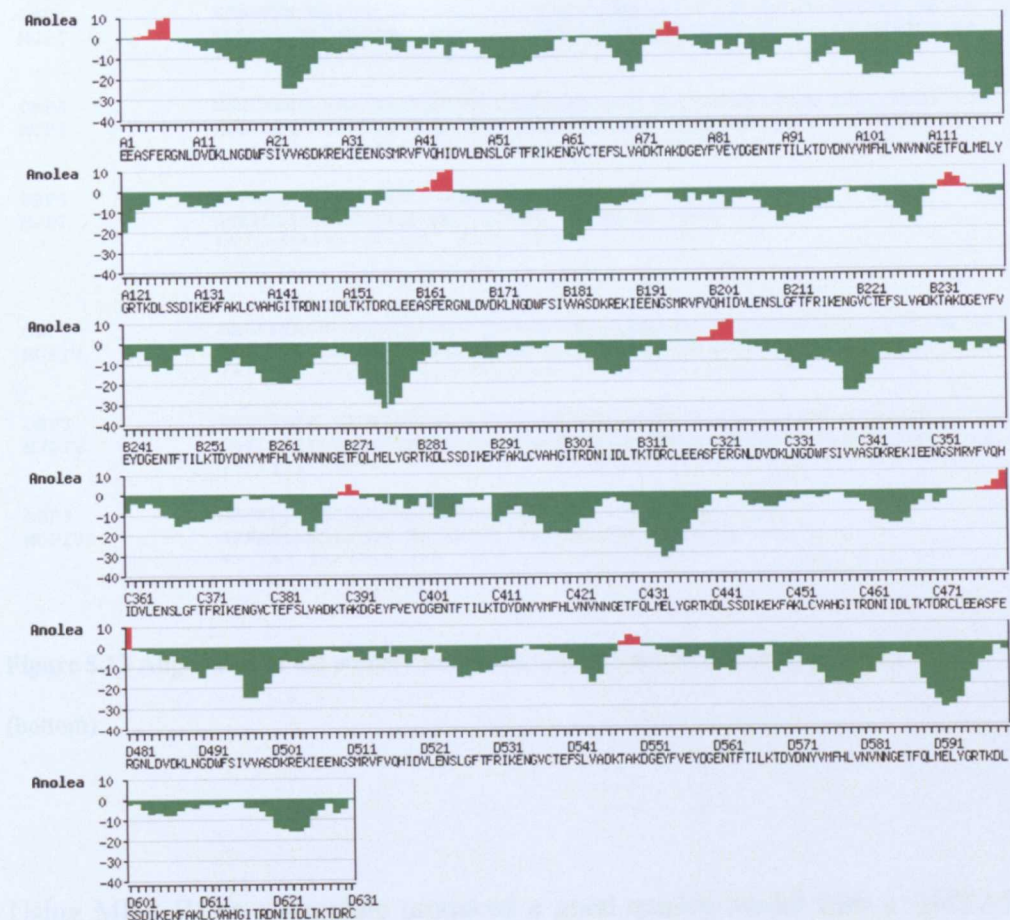


Figure 5.12 Assessment of the local quality of OBP3 modelled on A2U using ANOLEA. All four monomers are shown. Green indicates a negative, favourable, energies for a particular residue. Red indicates a positive, unfavourable, energy. Overall the local quality is good.

OBP3 has high sequence identity to a number of other molecules, notably MUP I and MUP IV (also known as nasal MUP). The crystal structures of these proteins were solved at higher resolution than A2U; 2.0 Å and 0.96 Å respectively (A2U was solved at 2.9 Å resolution) (Timm et al. 2001; Perez-Miller et al. 2010). The sequence alignments are shown in figure 5.13, MUPI has 62 % sequence identity to OBP3 whilst MUP IV has 70 % identity.

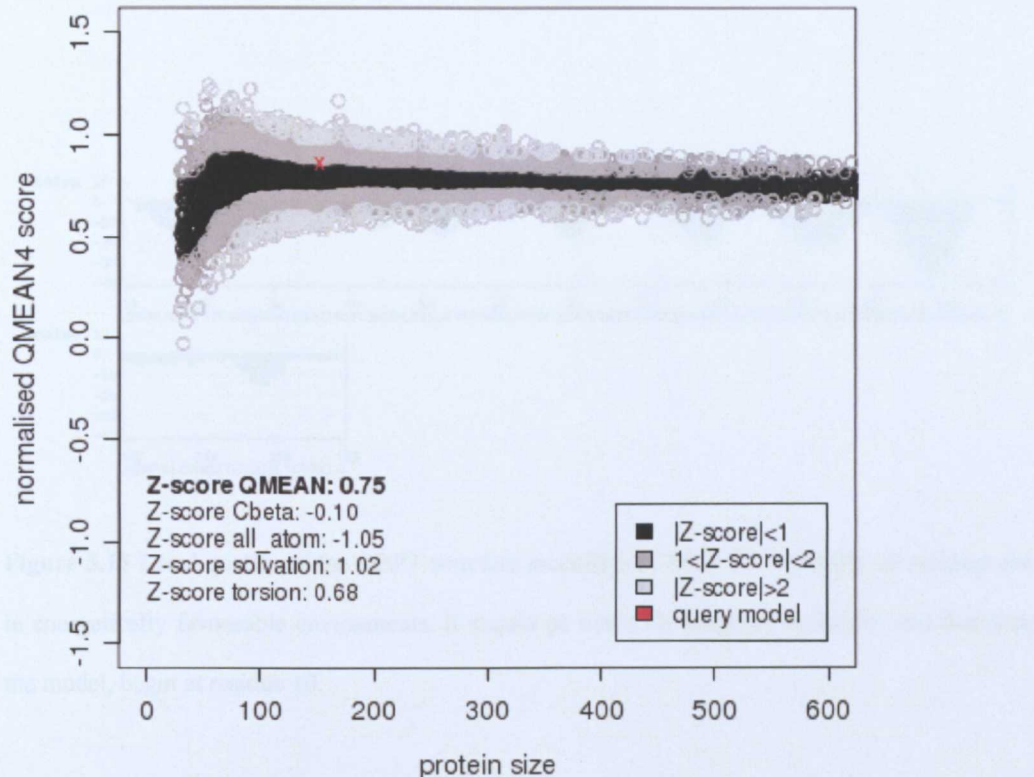


Figure 5.14 The QMEAN4 score of OBP3 modelled on MUP IV. The score (shown as a red cross) is much higher than average score across the reference set of PDB structures.

5.1.5. ANOLEA results

The results from ANOLEA (figure 5.15) suggested all the modelled residues were in favourable environments with most having strongly negative energies, the exception was lysine 124 which had a slightly positive energy. It is important to note that the crystal structure of MUP IV starts at residue 10, so therefore the residues that were positive at the N-terminus of the A2U modelled structure were not present in this model. This may all be reflective of the N-terminus being quite flexible and therefore there isn't sufficient electron density to describe this part of the molecule. In the NMR spectra of OBP3 all of the residues are observable with

the exception of the N-terminal histidine-tag and linker. Therefore, ideally, this section of the molecule needed to be modelled.

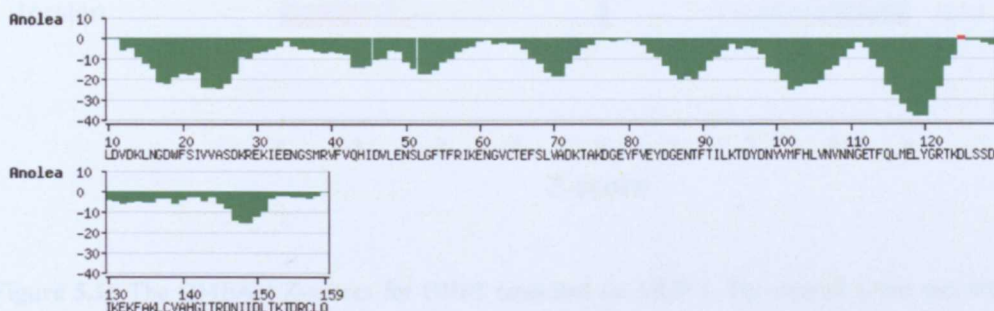


Figure 5.15 Local quality of the OBP3 structure modelled on MUP IV. Virtually all residues are in energetically favourable environments. It should be noted the template molecule, and therefore the model, begin at residue 10.

The crystal structure of MUP I is reported to be un-liganded and therefore was useful to use as a model template despite having the lowest sequence identity. The QMEAN score was 0.8 so compared to using the other templates the model is similarly robust. The QMEAN Z-score was 0.06, which is lower than for the other models, however it is still above the average (0) for the reference set of high resolution structures. As can be seen from figure 5.16 only the all atom interaction score was below zero, the other Z-scores were positive



Figure 5.16 The QMEAN Z-scores for OBP3 modelled on MUP I. The overall score was 0.06, however all the Z-scores apart from the all-atom interactions were above 0.

On a local basis, the energies for most amino acids were negative (figure 5.17). There was again a set of positive energies at the N-terminus and five other residues that have slightly positive energies (Glu 60, Lys 76, Asp 77, Ile 92 and Lys 94).

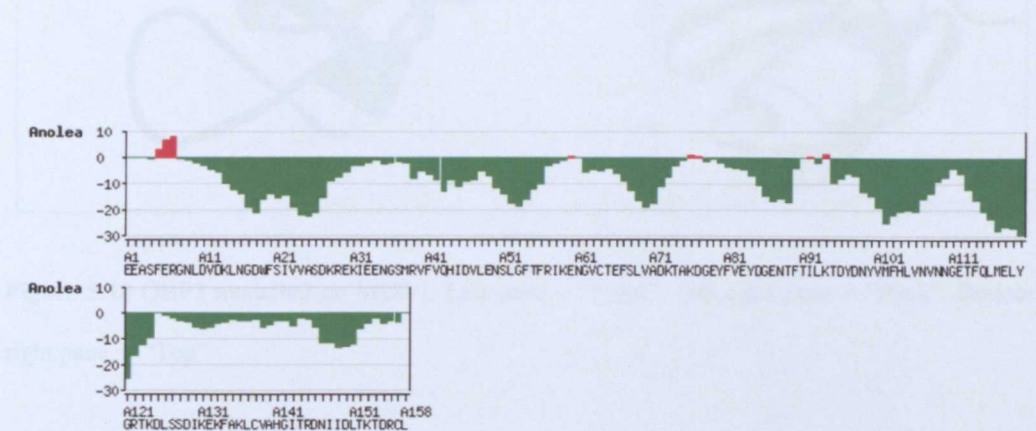


Figure 5.17 ANOLEA results for OBP3 modelled on MUP I.

The model of OBP3 using MUP I as a template is shown in figure 5.18. Again it possesses a typical lipocalin fold, the barrel is capped at the “bottom” by the N-terminal 3_{10} helix and a short β -strand. The C terminal α helix is packed against the face of the β -barrel. The top of the barrel contains the opening into the cavity. Multiple small cavities are detected, but a ligand binding pocket, like that described for A2U is not detected. There is no obvious ligand entry portal, leaving the cavity closed to external solvent.

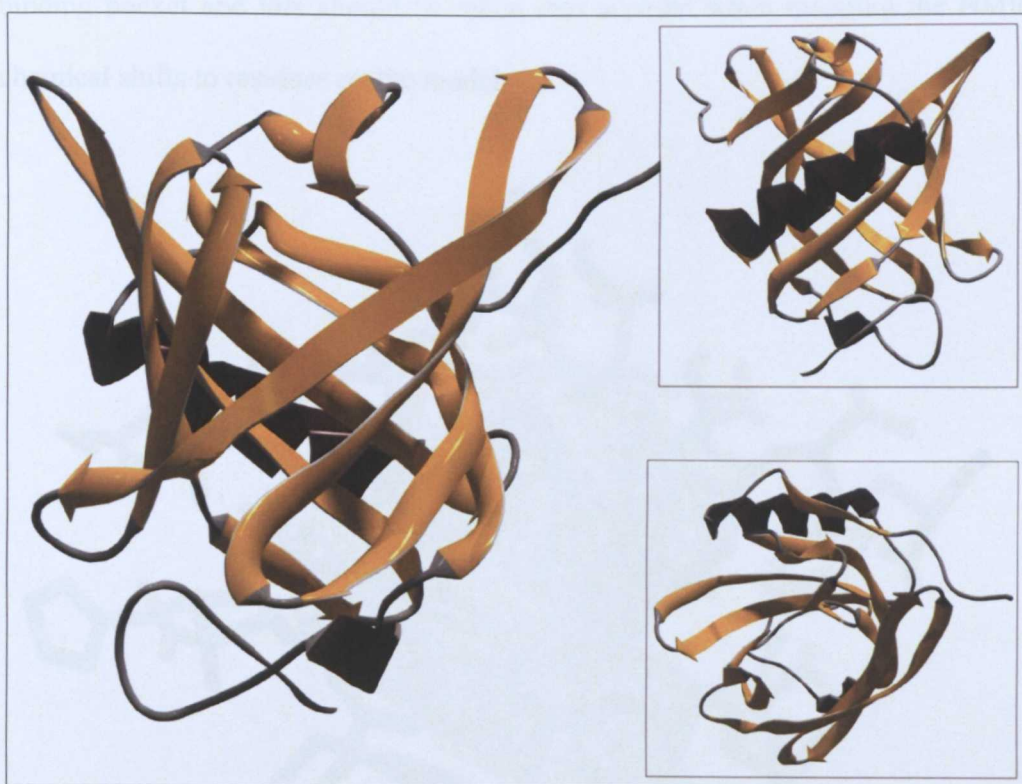


Figure 5.18 OBP3 modelled on MUP I. Left pane = “Front”. Top right pane = “Back”. Bottom right pane = “Top”.

Modelling OBP3 on MUP 1 bound to Sec-butyl thiazoline (1I06) however reveals a continuous binding pocket (shown in figure 5.9). This model of OBP3 had a high QMEAN4 score of 0.85, which was close to the structure modelled on MUP IV.

The QMEAN Z-score was also high, at 0.7, so it appeared overall the quality of the model was good. This binding pocket contains all the residues that line the pocket of A2U (or the equivalent residues) (Chaudhuri et al., 1999), plus a number of additional residues. The pocket seems to match the shape of Sec-butyl thiazoline (figure 5.19), so is unlikely to be the apo-form of the binding pocket. Like the structure modelled on apo-MUP I, there were also a number of smaller cavities. Overall it appears difficult to model the exact residues that line the ligand binding pocket and this should be taken into account when mapping the NMR chemical shifts to residues on the model.

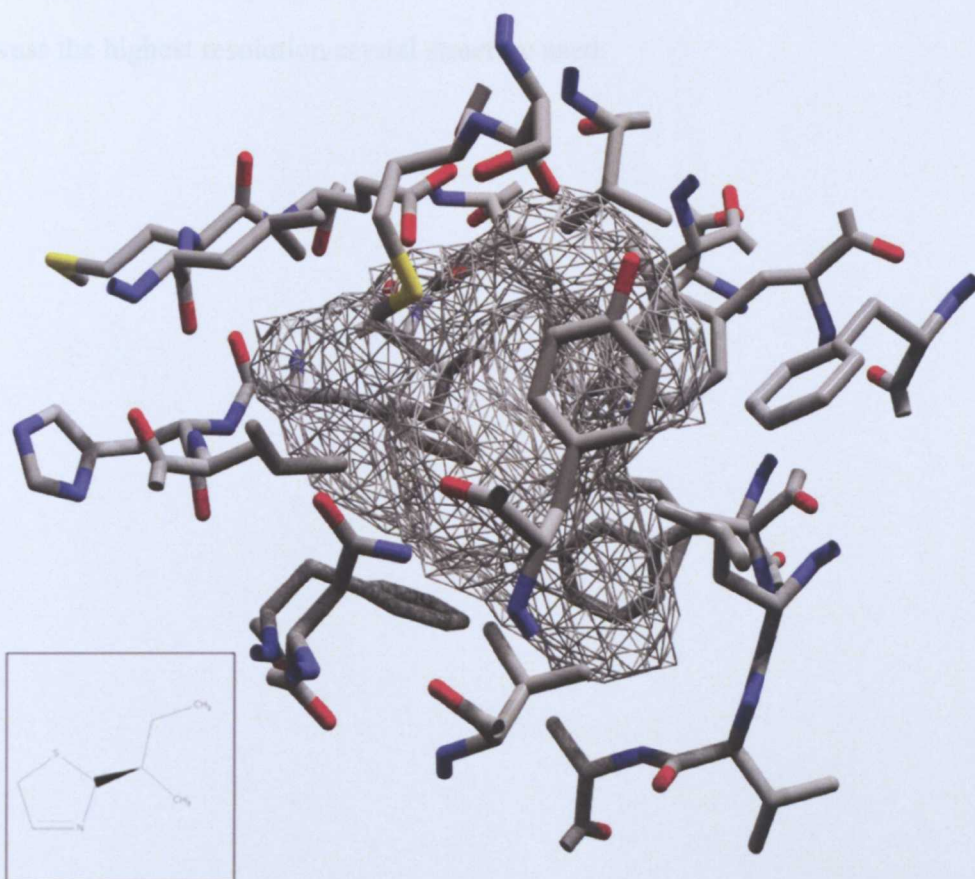


Figure 5.19 The pocket of OBP3 modelled on MUP I bound to Sec butyl thiazoline (shown in box). The pocket volume is 428 \AA^3 .

Overall the 5 models of OBP3 constructed, unsurprisingly, gave a similar fold. Figure 5.20 shows the 3 proteins used as models (A2U, MUP I and MUP IV) as well as OBP3 aligned by their primary sequences. The sequences are highlighted to denote their secondary structures (in the case of OBP3 the final secondary structure modelled is shown). This alignment demonstrates that despite differences in the exact residues each protein contains there is a general consensus. The majority of β -strands start within a residue of one another. The OBP3 shows some additional β -strand character in residues between β A and β B, which may be a result of over optimisation and may not actually be present. It is interesting to note how many extra 3^{10} helices MUP IV has, particularly as this was the highest resolution crystal structure used.

MUIV	EEATSKQNINVEKINGEWF	SILLASDKBEK	EEHGSMRV	FVEHIHVLENSLAFKFTVI	60
MUPI	EEASSTGRNFNVEKINGEWH	TIILASDKREK	IEDNGNFRILF	LEQIHVLEKSLVKFHTVR	60
OBP3	EEASFERNLDVD	KINGDWF	SIVVASIKREK	EENGS	60
A2U	EEASSTRGNLD	VAKINGDWF	SIVVASIKREK	EENGS	60
	***:	*:			
MUIV	DGECSEIFLVADK	TEKAGEYSVM	YDGFNTFTIL	KTDYDNIIMFHLIN	120
MUPI	DEECSELSMVADK	TEKAGEYSVT	YDGFNTFTIP	KTDYDNFLMAHLIN	120
OBP3	NGVCTEFSLVADK	TAKDGEYFVEYD	GENTFTIL	KTDYDNYVMFHLVNVN	120
A2U	NGECRELYLVAYKT	PEDEGEYFVEYD	GNTFTIL	KTDYDRYVMEHLIN	120
	:	*:			
MUIV	GRKADLN	SDIKEKFVKLCEE	HGIIKENI	IDLTKTNRC	162
MUPI	GRPDLS	SDIKERFAQLCEE	HGIIKENI	ILSNANRCLQARE	162
OBP3	GRTKDL	SSDIKEKFAKLCVA	HGITRDNI	IDLTKTDRCLQA--	160
A2U	GRTKDL	SSDIKEKFAKLCVA	HGITRDNI	IDLTKTDRCLQARG	162
	**	*****:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:			

Figure 5.20 Alignment of the primary sequences of the proteins used to model the OBP3 structure on. Secondary structure is indicated by green = β -strand, purple = α -helix and red = 3^{10} helix. The secondary structure shown for OBP3 is the final model used.

A2U is unusual in the fact that the structure is tetrameric, and due to the lack of evidence that OBP3 is at all multimeric (particularly in solution where NMR experiments are recorded), it is difficult to justify using this as a model. The highest quality structure was modelled on MUP IV, however identification of a ligand binding cavity proved impossible, and would make interpretation of NMR chemical shift maps somewhat complicated. This also applies to the model based on apo-MUP I, which appeared to have a number of disconnected small cavities. The structure of MUP I bound to the ligand sec-butyl thiazoline is of high quality and contains a substantial binding pocket that can at least be used as a good guide to the residues that are in the binding pocket. Additionally this protein is bound to a ligand that is very similar in structure to 2-isobutylthiazole. This model was therefore used to map the NMR chemical shifts.

5.2.6 Binding of 2-isobutylthiazole to OBP3 by heteronuclear NMR spectroscopy

NMR chemical shifts are very sensitive to changes in the local environment. The $^1\text{H} / ^{15}\text{N}$ HSQC ligand titration monitors the change in chemical shift of the backbone NH atoms. With nearly full assignments of unbound and 2-isobutylthiazole bound OBP3, a titration to monitor formation of this complex was straightforward. A 1 mM sample of ^{15}N labelled OBP3 was titrated with 2-isobutylthiazole (IBT) in 0.1 mM increments until there was no further peak movement, a $^1\text{H} / ^{15}\text{N}$ TROSY experiment was recorded at each increment and the spectra overlaid to track peak movement.

Overall the chemical shift perturbations (CSPs) of the peaks of OBP3 upon binding to IBT were relatively small, this is shown in figure 5.21 which shows the start of the titration (unbound OBP3, in blue) and the end of the titration (OBP3 fully saturated by IBT, in red). Most peaks did not move more than 1 to 2 ppm. In some cases there are some peaks that are only present when bound to 2-isobutylthiazole and some peaks that are unassigned, as described in chapter 4. It may be the case that these residues are highly shifted upon binding, but without assignments of both the unbound and bound forms this can not be commented on fully.

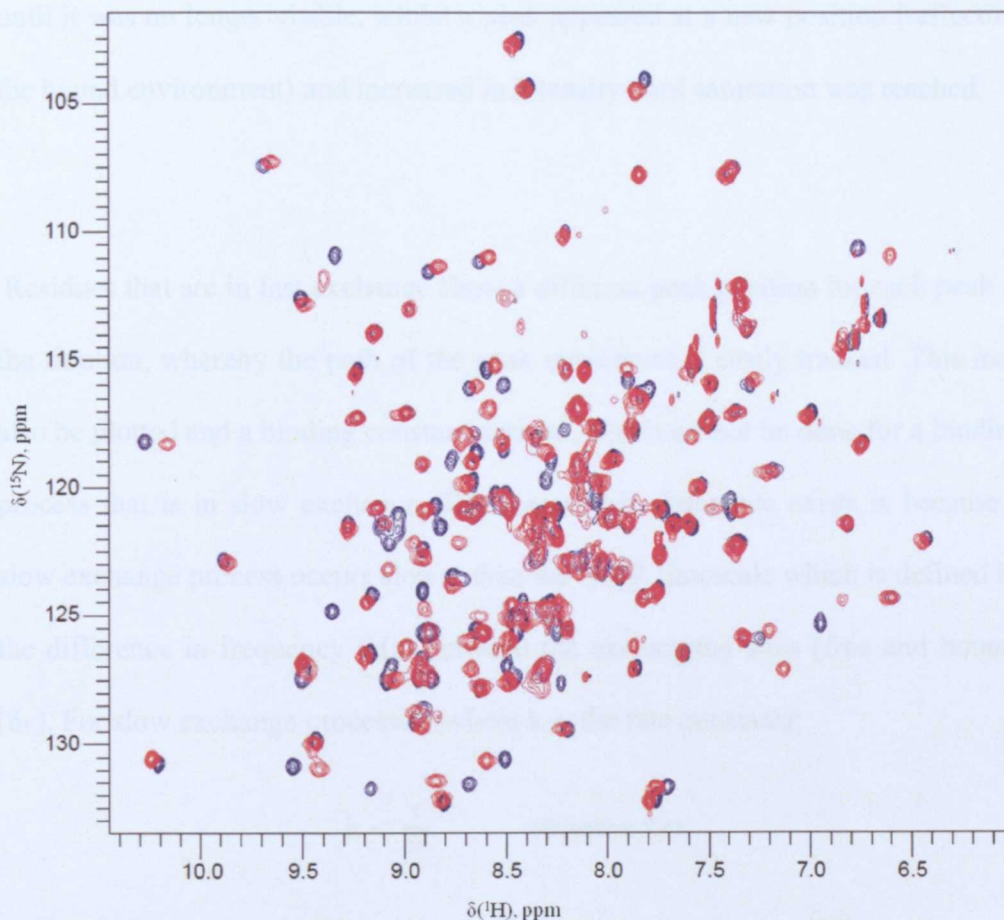


Figure 5.21 The unbound (blue) and bound to 2-isobutylthiazole (red) $^1\text{H} / ^{15}\text{N}$ TROSY spectra of OBP3. Overall the spectrum of bound form does not differ hugely from that of the unbound form shown by the small chemical shift perturbations.

The movement of peaks through the titration can be used to study the mode of binding. The movement seen in chemical shifts in the NMR experiments shows chemical exchange is taking place.. Figure 5.22 shows the movement of two peaks, selected as they are representative of the movement of the majority of peaks. It was found that most residues were in slow exchange. This conclusion was reached because as saturation took place the intensity of the peak in the starting position (reflecting the unbound environment of the residue) decreased

until it was no longer visible, whilst a peak appeared at a new position (reflecting the bound environment) and increased in intensity until saturation was reached.

Residues that are in fast exchange show a different peak position for each peak in the titration, whereby the path of the peak movement is easily tracked. This may also be plotted and a binding constant derived, which cannot be done for a binding process that is in slow exchange. The reason this difference exists is because a slow exchange process occurs slower than the NMR timescale which is defined by the difference in frequency (Hz) between the exchanging sites (free and bound) ($\delta\nu$). For slow exchange processes (where k is the rate constant);

$$k < \frac{1}{\delta\nu} \quad (\text{Equation 5.4})$$

This leads to resonances being visible for both the unbound and bound forms at the same time. Fast exchange processes occur more quickly than NMR timescale and so an averaged resonance is recorded at each increment.

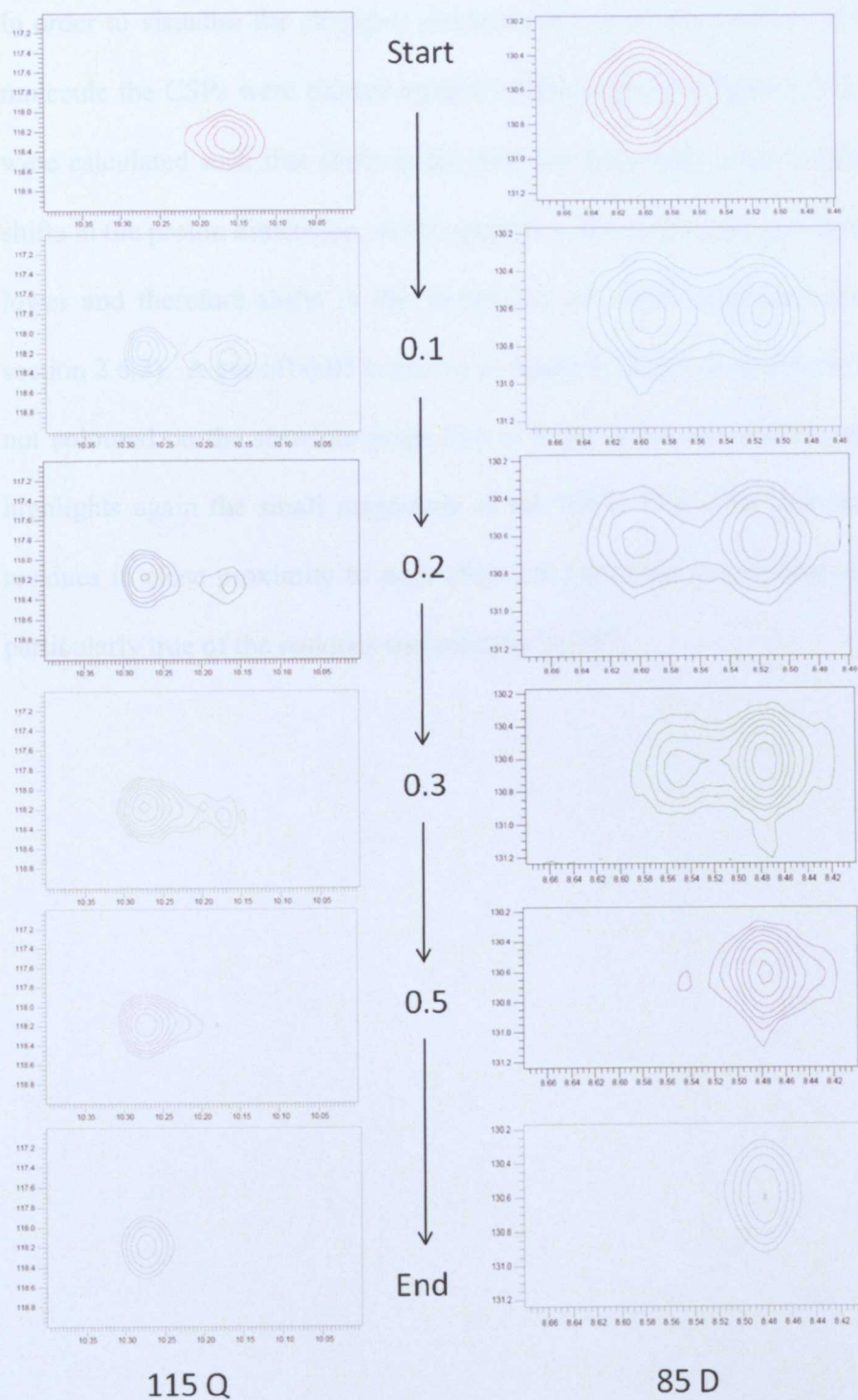


Figure 5.22 Residues 115 Q and 85 D are representative of most residues that change environment, as they undergo slow exchange from the unbound (Start) to the bound (End) form. The values correspond to the millimolar concentration of 2-isobutylthiazole titrated into a 1 mM sample of OBP3.

In order to visualise the changing environments of all the residues of the protein molecule the CSPs were plotted against residue number (Figure 5.23). The CSPs were calculated such that shifts in the nitrogen dimension were weighted against shifts in the proton dimension, as the spectral width of the nitrogen dimension was lower and therefore shifts in this dimension are more significant (described in section 2.6.3). A cut-off 0.05 is shown in figure 5.23, below which the shifts were not coloured on the structure maps due to being much less significant. The plot highlights again the small magnitude of the CSPs. It is clear that small sets of residues in close proximity to each other are perturbed in a similar manner, this particularly true of the residues surrounding Val 40.

The CSPs were then plotted on the homology of model of OBP3 (modelled on MUP1:1I06). This is shown in figure 5.24, where the perturbed residues are displayed on a ribbon diagram of OBP3 as space filling backbone atoms. The most highly perturbed residues are the darkest blue, decreasing to the lightest blue for the least perturbed residues.

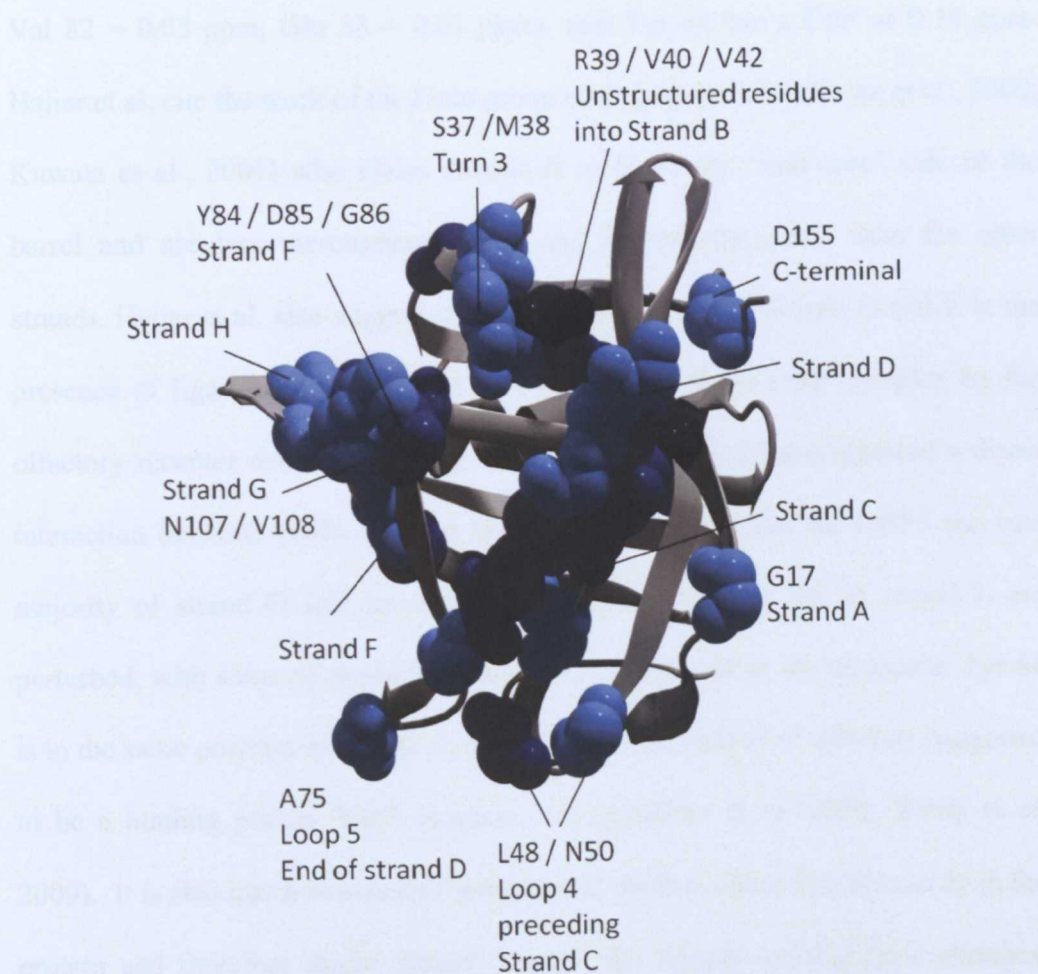


Figure 5.24 Amide proton groups showing chemical shift perturbations of OBP3 upon binding to 2-isobutylthiazole that were greater than 0.05 ppm. The darkest blue residues are the most perturbed.

From this it is clear that β -strands C and D have a large number of perturbed residues, with some residues that are quite strongly perturbed. This is interesting

as, during their molecular dynamic simulations, Hajjar et al. (2006) describe the transient opening of a gap between sections of β -strands D and E of OBP1 associated with the presence of ligand, whilst the structural hydrogen bonding of the rest of the molecule remained intact. Phe 80 and Tyr 81 are unassigned for OBP3, whilst the rest of strand E is only slightly perturbed (Glu 79 = 0.05 ppm, Val 82 = 0.03 ppm, Glu 83 = 0.01 ppm), and Tyr 84 has a CSP of 0.14 ppm. Hajjar et al. cite the work of the Goto group on β -lactoglobulin (Forge et al., 2000; Kuwata et al., 2001) who claim strands B to E are the “non-core” side of the barrel and are less thermodynamically and kinetically stable than the other strands. Hajjar et al. also suggest that the opening between strands D and E in the presence of ligand maybe required for recognition of the OBP complex by the olfactory receptor neurons (ORNs). However few studies have reported a direct interaction between OBPs and the ORNs. It does seem that for OBP3 the vast majority of strand D and some residues, particularly Tyr 84, of strand E are perturbed, with some of the higher perturbations seen across the molecule. Tyr 84 is in the same position as Tyr 82 in rat OBP1 and porcine OBP which is suggested to be a binding pocket “cap” (section 1.4) (Meillour et al. 2009; White et al. 2009). It is also true however that a number of these residues line the cavity in the protein and therefore might directly contact the ligand, causing their chemical shift perturbation. Strands F, G, and H also contain a small number of perturbed residues. The largest perturbations are seen in the residues preceding, and at the very start of, β -strand B. This is one of a number of unstructured and loop regions that appear to be involved in binding, also including loops 4, 5 and 6. Turn 3 is also perturbed upon binding, this forms part of a number of perturbed residues surrounding the opening at the top of the cavity, where ligand entry is thought to

take place. The N and C termini of the protein do not seem to change environments upon binding; Asp 155 at the C-terminal is perturbed (0.09 ppm), whilst Gly 17, which is after the 3¹⁰ helix at the N-terminus was affected to similar degree (0.07 ppm). Although from x-ray crystal structures of the similar proteins used for modelling, the N-terminus seemed to be flexible, this lack of change in shift suggests the terminus is not fixed upon binding. This could possibly indicate that it remains somewhat flexible and the NMR peaks are an average position of these residues.

It is difficult to see from figure 5.24 which of the perturbed residues also line the cavity of the protein, and are possibly in contact with the ligand. Figure 5.25 has the same colouring as figure 5.24 but shows only the largest cavity in the molecule (428 Å³), viewed from the “top” of the protein looking down into the cavity.

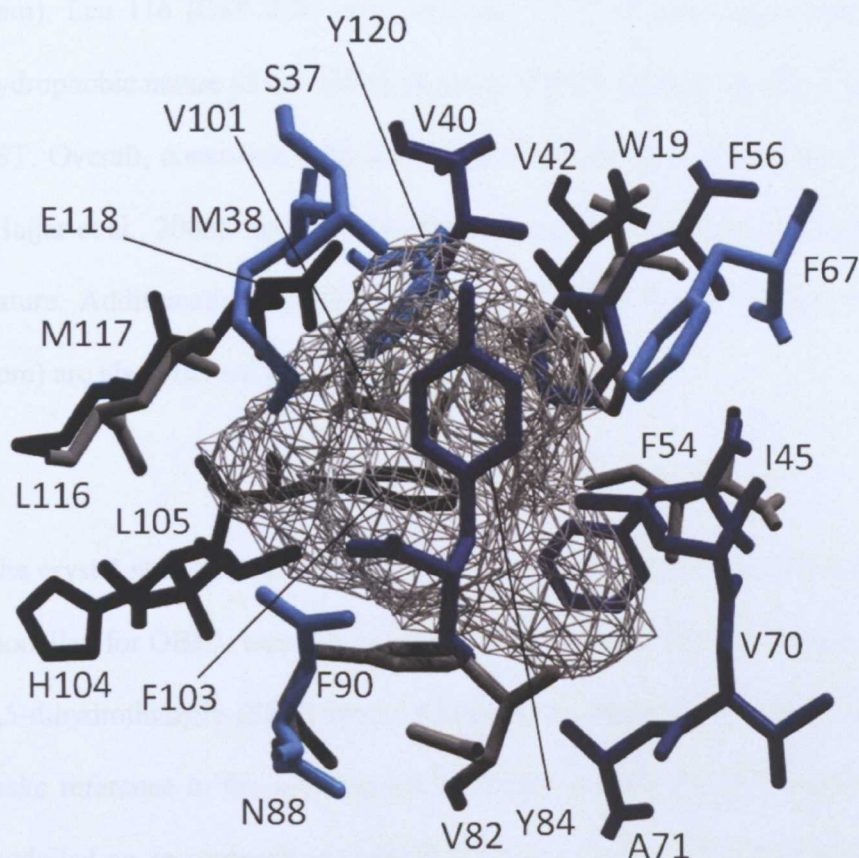


Figure 5.24 The binding cavity of the homology model of OBP3. Residues that are unassigned are shown in black, residues perturbed less than 0.05 ppm are grey whilst those perturbed are coloured from light to dark blue (as in figure 5.23) as the CSPs increase.

The most perturbed residue of all, Arg 39, is not in the cavity, however it does appear to be in a smaller cavity (42 \AA^3) in close proximity to the large cavity. Residues that were not assigned (in either the unbound or bound forms) are shown in black (Phe 103 to Leu 105, Met 117 and Glu 118). Interestingly not all residues that line the cavity appear to significantly change environment upon binding to IBT, notably Phe 90 (CSP = 0.03 ppm) and Trp 19 (0.02 ppm), despite having large side chains containing rings. Val 82 is not significantly perturbed (0.03 ppm) whilst Val 42 is highly affected (0.32 ppm) and, to a lesser extent, Val 40 (0.18

ppm). Leu 116 (CSP 0.04 ppm) and Met 38 (0.07 ppm) also contribute to the hydrophobic nature of the pocket but only Met 38 seems to be involved in binding IBT. Overall, consistent with the nature of the binding pocket and other studies (Hajjar et al., 2006), the majority of the the perturbed residues are hydrophobic in nature. Additionally Arg 39 (0.33 ppm), Ser 37(0.08 ppm), and Asn 88 (0.07 ppm) are also affected by the binding of IBT.

The crystal structure of MUP IV, which has a very similar binding pocket to that modelled for OBP3, was solved bound to the male mouse pheromone 2-sec-butyl-4,5-dihydrothiazole (SBT) (Perez-Miller et al., 2010). The authors of this paper make reference to the inverted cavity shape of MUP I (1I06) which OBP3 was modelled on, in comparison to MUP IV. Perez-Miller et al. found direct hydrogen bonding between side chain of Glu 118 and the nitrogen of SBT, with an additional hydrogen bond between the hydroxyl of Tyr 120 and the side chain of Glu 118. There are no other direct bonds reported, although Ser 21, Met 38, Val 40, Phe 54, Val 82 and Phe 103 are all reported to surround the ligand (Figure 5.26). This matches well with the perturbations found for OBP3 binding to 2-isobutylthiazole, though OBP3 appears to have a number of additional hydrophobic residues perturbed. Additionally Perez-Miller et al. make no reference to Asn 88 and Arg 39 being contacted, although they are more highly perturbed than most residues in OBP3 binding to IBT.

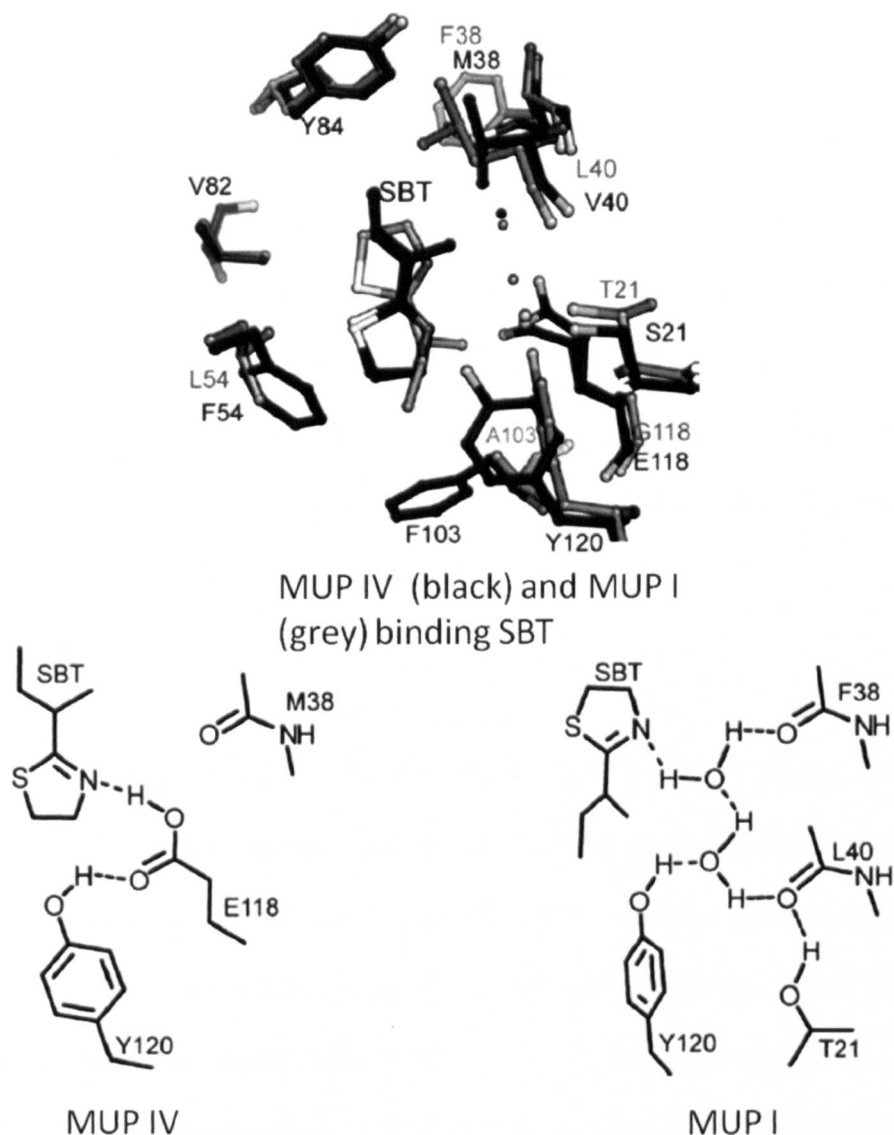


Figure 5.25 Residues in the binding pockets of MUP IV (black) and MUP I (grey) surrounding the ligand 2-sec-butyl-4,5-dihydrothiazole (Top). The crystal structures show the ligand binding in reversed orientations with altered hydrogen bonding. Adapted from (Perez-Miller et al., 2010)

5.3 Conclusions

In this chapter previous binding studies on rat OBPs were reviewed and used to select three sets of odours to study binding to OBP3 by ITC. ITC experiments

required careful planning to reduce the chance of plastics and environment odours binding to OBP3. It was found that the thiazoles (IBT and DMT) tested had higher affinities than had been found previously (Löbel 2001, 2002). Saturation with octanol, followed by extensive dialysis to remove the octanol had been used to remove ligands already bound at the end of purification (Löbel 2001). The experiments in this study revealed that binding to octanol was much weaker than to heterocyclic compounds and it may be that case that this ligand is not capable of displacing all the endogenous bound ligands. This would also account for the difference in binding stoichiometries achieved.

OBP3 modelled on MUP I bound to SBT (1I06) was used to map the resulting chemical shift perturbations resulting from NMR titrations of OBP3 with IBT. A number of residues in the binding pocket were found to be significantly perturbed, this included, but was not limited to, hydrophobic residues. The CSPs did not point to a clear reason why OBP3 had such high affinity for the thiazoles investigated. The presence of possible hydrogen bonds from Glu 118 or Tyr 120 could explain this, however Glu 118 was not assigned in the unbound form of OBP3, whilst there was only moderate perturbation of Tyr 120 (0.10 ppm). In addition the odours containing a double bonded oxygen or hydroxyl are also capable of such interactions. It may also be the case that the hydrophobicity of IBT (as judged by log P) is particularly well matched to the hydrophobic nature of the OBP pocket, although IBT is unremarkable compared to the other ligands tested by ITC with regard to its hydrophobicity or water solubility. The size of the binding pocket of OBP3 could also contribute to the high affinity interaction with

IBT, with the ligand being small enough to fit in the pocket, but big enough to make substantial contacts with the residues lining the pocket.

Finally, the α -helix of OBP3 doesn't contain any residues affected by the titration, suggesting it plays no part in binding. This is consistent with the fact that the α -helix has not been reported to be involved in the ligand binding of any other lipocalin.

6. USING DEFINED LIGAND SERIES TO EXPLORE THE BINDING OF OBP3

6.1 Introduction

In chapter 5 a range of odours was examined and found to have different binding affinities and thermodynamic characteristics when binding to OBP3. The heterocyclic compounds investigated had particularly high affinity interactions with OBP3. To study this further the particular residues affected on OBP3 binding to 2-isobutylthiazole were found using NMR titrations. It was difficult however to make judgements about the relative contributions of size, hydrophobicity and positioning of functional groups to the overall thermodynamics of binding. In this chapter a set of related ligands that bind to OBP3, each with a single change in structure were used to make it possible to make correlations and spot trends. The flexibility of the residues of OBP3 across this series of complexes was observed using heteronuclear NOE experiments. In addition, a set of isomers was used to look at the effect of ligand shape and orientation on binding affinity and thermodynamics. In order to judge how solvent accessible the residues of OBP3 were, in particular, the central cavity of OBP3, deuterium oxide (D₂O) exchange studies were conducted.

6.1.1 The γ -lactone series

Lactones, as described in chapter 5, are cyclic esters. They consist of a closed ring including a single oxygen atom. Lactones may be considered the result of a condensation of an alcohol group and a carboxylic acid group. The size of the

lactone ring is indicated by the Greek letter prefix (α -lactones are 3 member rings, β -lactones are 4 membered, γ -lactones are 5 membered and δ -lactones have 6 members). The rest of the name is derived from the precursor acid molecule (for example butyrolactone has 4 carbon atoms, valerolactone has 5 carbon atoms). γ -lactones and δ -lactones are the most stable, as the ring size minimizes the strain of the bond angles. α -lactones are only transiently detected, whilst β -lactones are only formed by synthesis using specific methods (Kramer et al., 2006).

The γ -lactones are particularly important in beverage production and the alkyl substituted γ -lactones used in this chapter have been identified as important sensory components in wine (Brown, 2007). The series of five γ -lactones used are shown in figure 6.1.

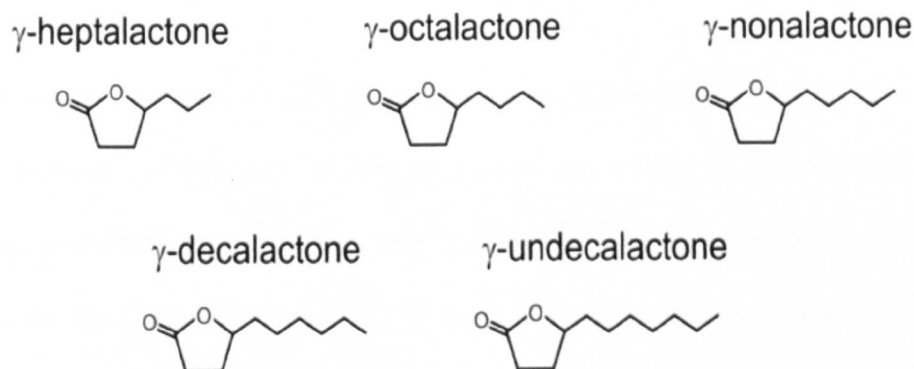


Figure 6.1 The γ -lactones used in this chapter. A set of alkyl substituted molecules which form a series, differing only in the addition of a CH_2 at each stage in the series.

Despite being highly structurally related (differing only in the addition of a CH₂ group), each of the γ -lactones has different and quite distinguishable, sensory attributes, which are listed in table 6.1. Overall the basic odour of the series is described as “fatty” however they also possess fruity odours which leads to their importance in the food and beverage industry.

Lactone	Molecular Weight	Odour Description
γ -heptalactone	128.17	Coconut / Nutty / Caramel
γ -octalactone	142.20	Coconut / Fruit / Raspberry
γ -nonalactone	156.22	Sweet / Honey / Vanilla
γ -decalactone	170.25	Peach / Apricot
γ -undecalactone	184.28	Fruity / Pear / Tropical

Table 6.1 Sensory attributes and molecular details of the γ -lactones used.

Increasing the γ -lactone chain length leads to an increase in hydrophobicity, a useful way of measuring this that is commonly used in the food industry is Log P. What Log P represents was described in section 5.1.2. Using calculated values of Log P for the γ -lactones gives a clear linear relationship. For the purpose of comparisons in this study the experimentally determined Log P value for IBT is plotted on the graph, indicating is more hydrophobic than γ -heptalactone to γ -nonalactone.

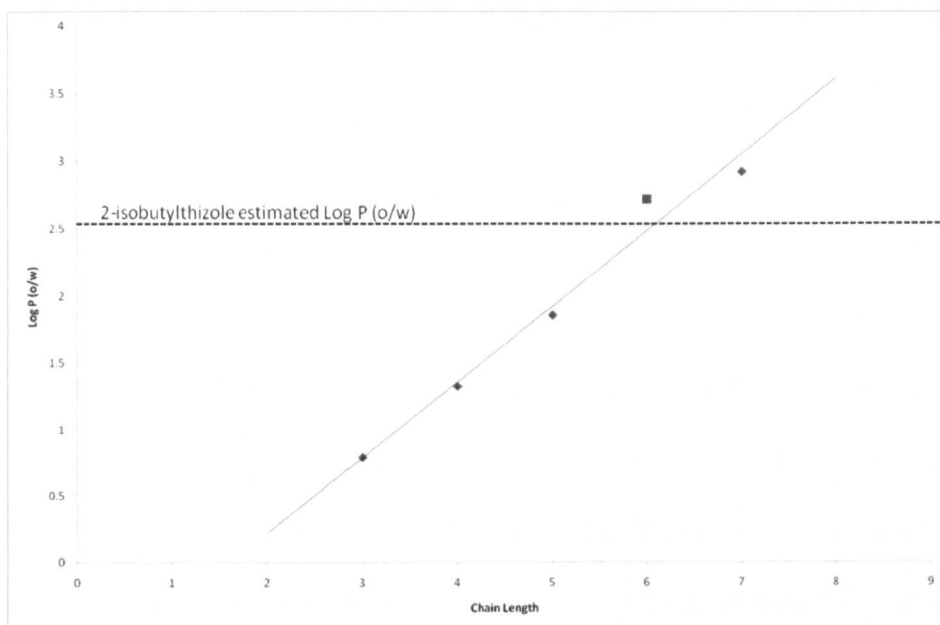


Figure 6.2 Log P values of the γ -lactones plotted using hydrocarbon chain length. The log P value for IBT is shown by a dashed line (Rabe, 2004)

6.2 Results and Discussion

6.2.1 D₂O exchange

The method used for hydrogen / deuterium exchange is described in section 2.6.3.5. Briefly, lyophilised OBP3 protein was solubilised in deuterium oxide and immediately submitted to $^1\text{H} / ^{15}\text{N}$ TROSY experiments. Initially experiments were run every 20 min (for 720 min) and then subsequently every 30 min for a further 60 hours. A final experiment was run after 7 days to confirm no further exchange had taken place. The reduction and eventual disappearance of the NH signals as they became ND and therefore were no longer observable by NMR was monitored by recording multiple $^1\text{H} / ^{15}\text{N}$ TROSY experiments.

Figure 6.3 shows the residues (with backbone atoms highlighted in green on the ribbon diagram) that had exchanged before the first experiment had been run (approximately 480 s). Three glycine residues in loop regions exchanged immediately; Gly 8 near the N-terminus, Gly 34 the second and third part of β -strand A and Gly 62 in loop 5 between β -strands C and D. Two threonine residues near the C-terminus (Thr 152 and 154) also rapidly exchanged. Other residues that had exchanged before the first experiment, also found in or near loop regions included Arg 27, Thr 74, Ala 75, Asp 85, Ile 92, Tyr 97, and Asn 110. Alanine 75 is particularly interesting as this residue was also perturbed upon binding to 2-isobutylthiazole, suggesting this residue may be quite mobile. All of the loop regions are modelled as being solvent accessible and this may account for the rapid exchange. The beginning of β strand H (Gly 111 to Thr 113), Glu 83 (β -strand E) and Asp 107 (β -strand G) all exchanged rapidly despite being part of β -barrel strands where hydrogen bonds between the β -strands are likely to take longer to exchange. This may suggest a lack of hydrogen bonding here, a degree of flexibility or that these residues are solvent accessible. Leu 126 to Asp 129 at the start of the α -helix also exchanged rapidly, showing this end of the helix to be more solvent exposed and possibly more flexible than the rest of the α -helix. Interestingly Tyr 120 exchanged rapidly, this residue was shown by the homology model to be in the binding pocket, and in nasal MUP is directly involved in hydrogen bonding to 2-sec-butyl-4,5-dihydrothiazole (SBT) (Perez-Miller et al., 2010) (section 5.2.8). Other proteins that are homologous to OBP3 have central cavities that are quite occluded to solvent (Sharro et al. 2003; White et al. 2009; Wang et al. 2011), and therefore it may be expected that these residues would remain unexchanged. Tyr 120 is at the end of β -strand H, as mentioned in the

previous chapter, this side of the barrel is considered to be the more stable, “core”, side and therefore the exchange seen may indicate that some water does occupy the empty binding pocket and this is then able to hydrogen bond to Tyr 120. It should be noted, however, that none of the other residues in the remainder of the pocket exchanged rapidly.

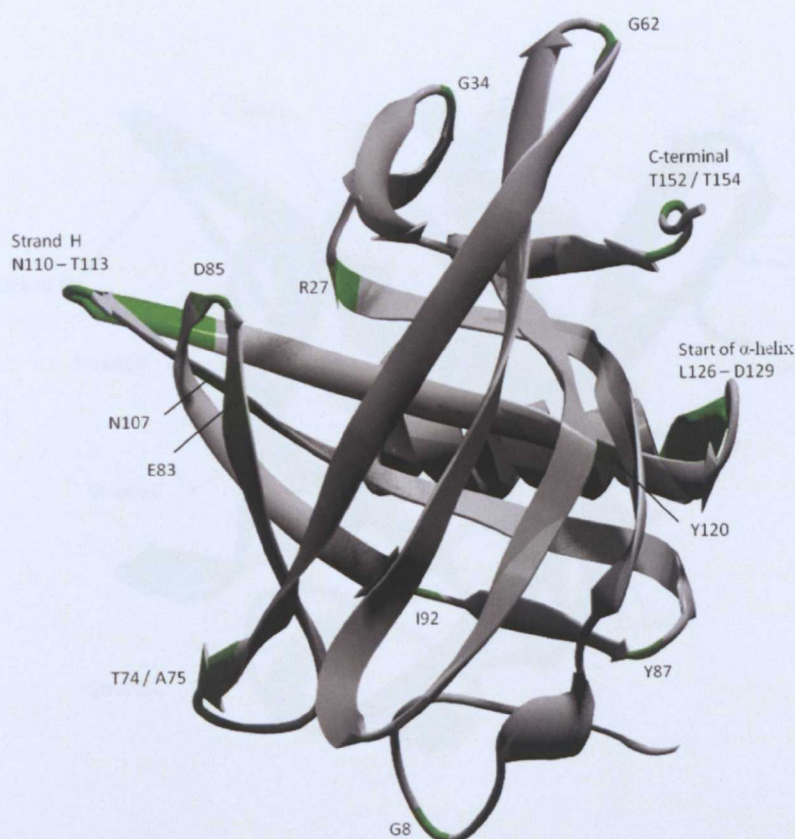


Figure 6.3 Residues rapidly exchanged with D₂O. The residues highlighted in green were not visible on the first ¹⁵N TROSY HSQC as they had already exchanged with the D₂O solvent

Figure 6.4 shows that the majority of β-strands C and D remained unexchanged after 72 hours. Hajjar et al. (2006) claim in the presence of ligand that strands D

and E are able to transiently open. Strand D in particular is clearly very stable, so the mechanism of strand opening would be interesting to observe. D₂O exchange experiments are not possible with a ligand bound to OBP3, as large volumes of D₂O would be required to solubilise the ligand, the alternative would be to bind the ligand to OBP3 prior to lyophilisation, however it is unlikely that enough of the ligand would remain bound to OBP3 due to its volatile nature.

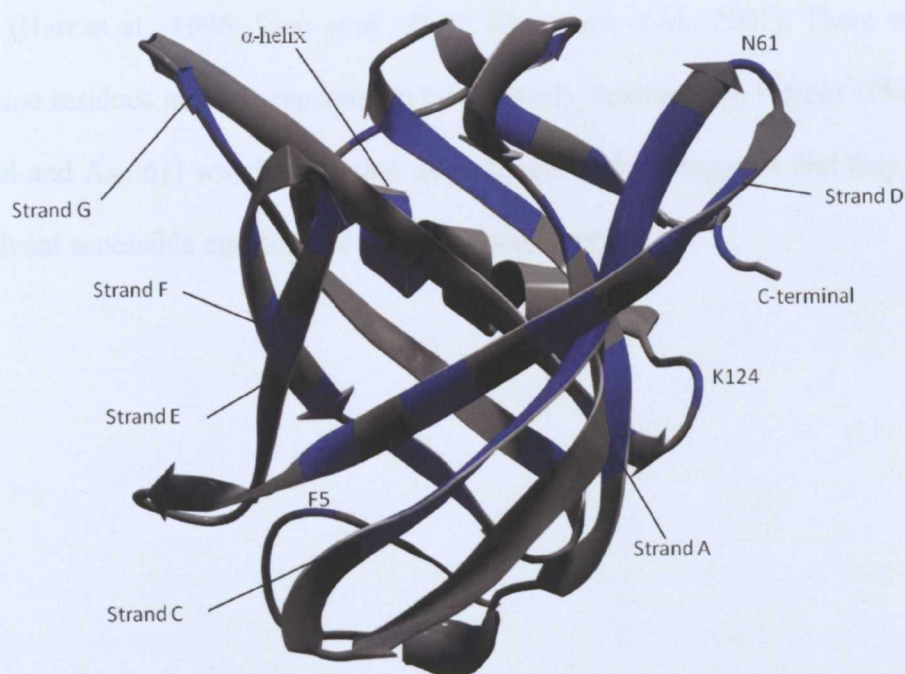


Figure 6.4 Residues that remained unexchanged after 72 hours. Unexchanged residues are highlighted in blue. A number of residues in the binding pocket and β -strands C and D remain unexchanged.

β -strands E, F and G also contained some residues that remained unexchanged after 72 hours, as did larger parts of β -strands A and B. This is likely to be due to their being part of stable inter-strand hydrogen bonds.

A large part of the α -helix had also not exchanged after 72 hours, suggesting it may be stably “pinned” to the β -barrel and not flexible as has been found in some β -barrel and β -clam proteins, where it forms a capping motif over the ligand pocket (Herr et al., 1996; Liou et al., 2002; Kouvatsos et al., 2007). There were also some residues in what appeared to be relatively flexible loop regions (Phe 5, Lys 124 and Asn 61) which were also unexchanged, which suggests that they are less solvent accessible and flexible than they may appear.

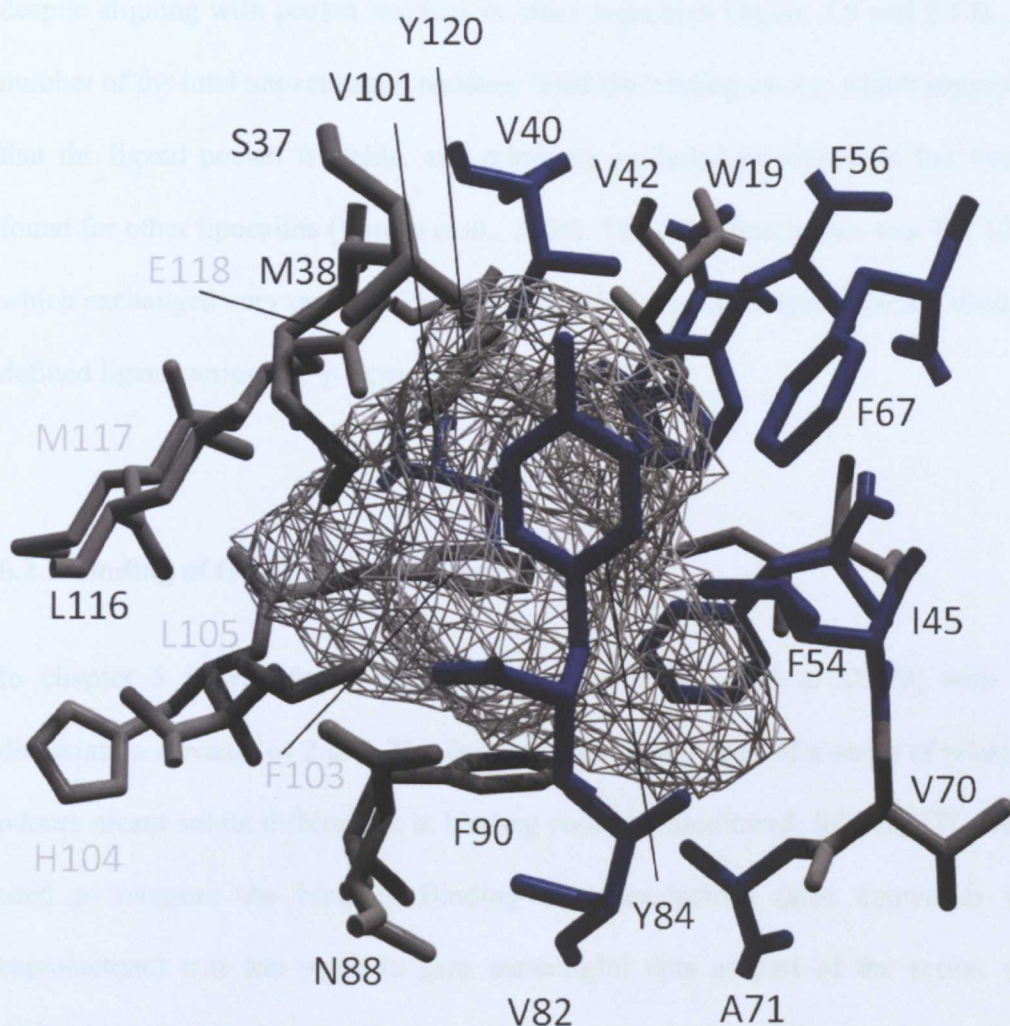


Figure 6.5 Unexchanged residues in the ligand binding pocket. Residues that did not exchange within 72 hours are highlighted in blue. Unassigned residues have their labels greyed out. The cavity is denoted by the hatched area. A number of aromatic residues in particular remain unexchanged.

Overall, half of the assigned residues in the cavity were still unexchanged after 72 hours (Figure 6.5), with many also being part of β -strands C or D. Phe 90 had exchanged within 72 hours, as had Asn 88, however, unlike Asn 88, Phe 90 was only minimally perturbed upon binding to 2-isobutylthiozole. This raises the question of whether or not Phe 90 is actually part of the ligand binding pocket,

despite aligning with pocket residues in other sequences (figure 5.9 and 5.13). A number of the total unexchanged residues lined the binding cavity; which suggests that the ligand pocket is stable, and relatively occluded to solvent as has been found for other lipocalins (Barratt et al., 2005). The exception to this was Tyr 120 which exchanged very rapidly. The binding pocket was investigated further using a defined ligand series, the γ -lactones.

6.2.2 Binding of OBP3 to the γ -lactones by ITC

In chapter 5 it was found that γ -decalactone bound well to OBP3, with a dissociation constant of 2 μ M. The fact that this odour is part of a series of related odours meant subtle differences in binding could be monitored. Initially ITC was used to measure the binding. Binding of γ -hexalactone (also known as γ -caprolactone) was too weak to give meaningful data as part of the series; γ -dodecalactone was also weakly bound so wasn't included. Other studies in this laboratory on OBP1 found no binding at all to γ -dodecalactone, so the fact the OBP3 did show some interaction may reflect the difference in the pocket size of OBP1 (326-328 \AA^3 (White et al., 2009)) and OBP3 (428 \AA^3 , homology model).

ITC was carried out using the same protocol as used previously for γ -decalactone in chapter 5. In figure 6.6 the dissociation constants are plotted against the length of the hydrocarbon side chain, not counting the lactone ring (for example, a chain length of 3 is γ -heptalactone).

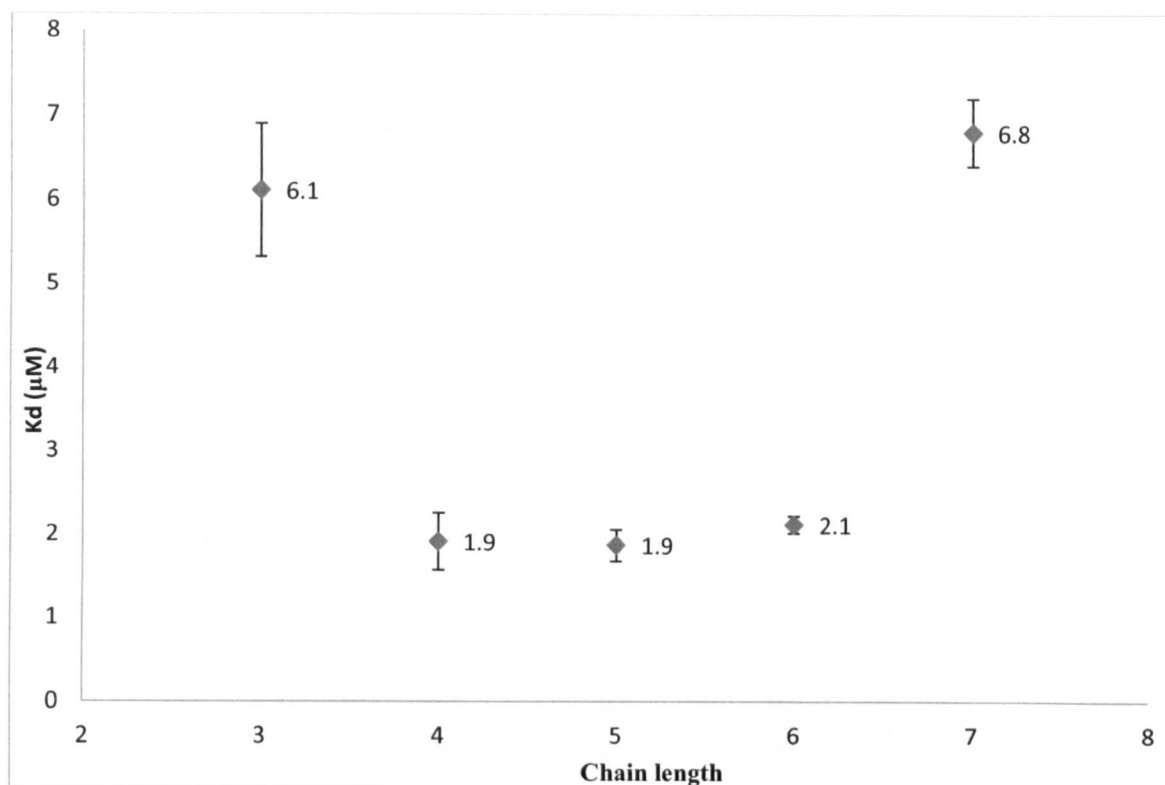


Figure 6.6 Dissociation constants of the γ -lactones binding to OBP3. Overall the affinities are quite similar, with the middle of the series showing affinities within error of each other, which are also the higher affinity interactions. Error bars represent the standard error of the mean of duplicate experiments.

For γ -octalactone, γ -nonalactone and γ -decalactone the dissociation constants are within error of each other ($K_d = \sim 2 \mu\text{M}$). The smallest and largest γ -lactones in the series have dissociation constants approximately three times weaker than the rest. This trend would support the fact that the extremes outside of the series are even more weakly bound and may suggest a “window” of optimal size in γ -lactone binding affinities that indicates that accommodating the hydrocarbon chain is a factor in OBP3 binding. However, overall the affinities of the series are

within the same range (low micromolar affinity) and therefore the ΔG° values reflect that the overall favourability of the interactions is similar.

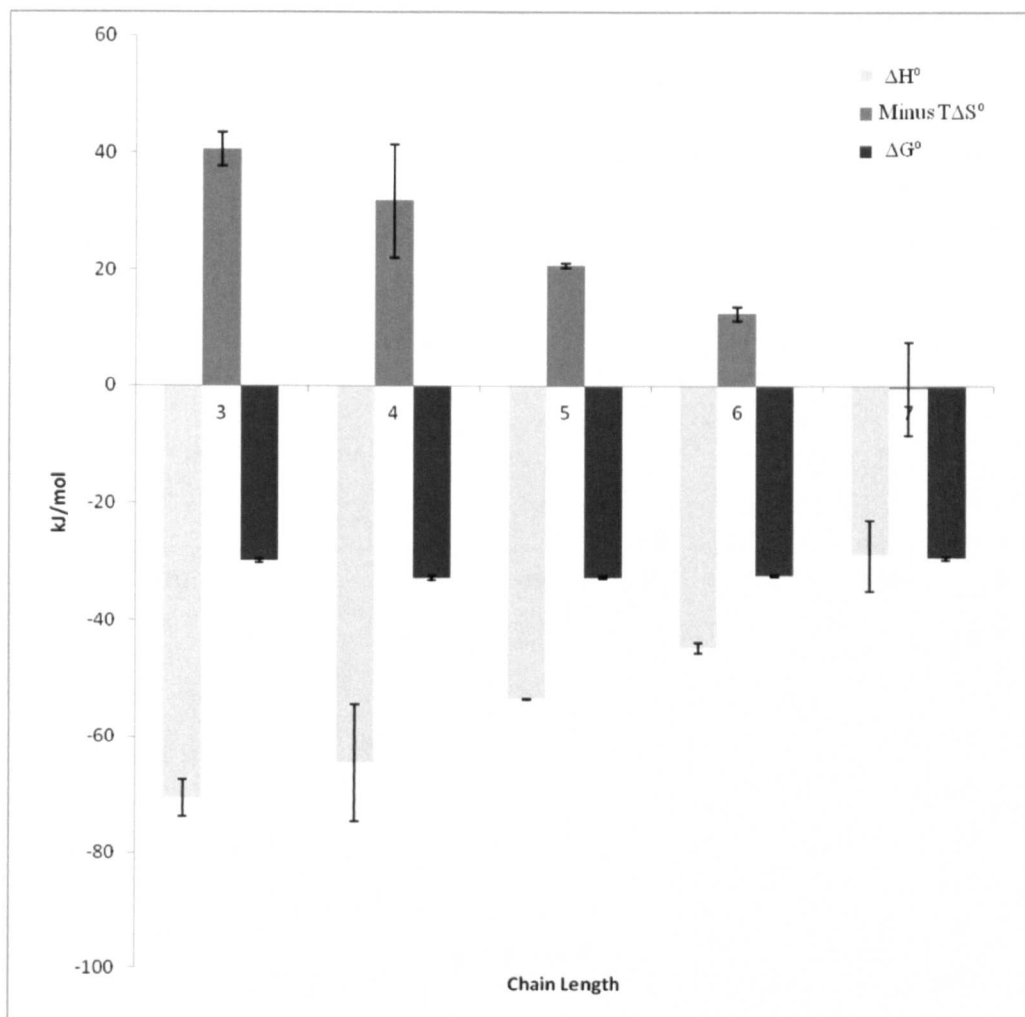


Figure 6.7 Thermodynamic parameters of γ -lactones binding to OBP3. The ΔG° values remain quite constant, although the ΔH° values become less favourable and the ΔS° values more favourable as chain length increases.

Figure 6.7 shows the parameters plotted against the hydrocarbon chain length of the γ -lactones. Minus $T\Delta S^\circ$ is plotted rather than $T\Delta S^\circ$ to make the bars easier to distinguish. A trend of a decreasing change in enthalpy was seen as the

hydrocarbon chain length increased, with the most negative enthalpy change seen for γ -heptalactone and the least negative change seen for γ -undecalactone. The entropy across the series became less negative (seen in figure 6.7 as a less positive minus $T\Delta S^\circ$). The entropy change for γ -undecalactone was almost zero, being in fact slightly positive. If $T\Delta S^\circ$ is plotted against ΔH° a strong linear relationship is seen (figure 6.8).

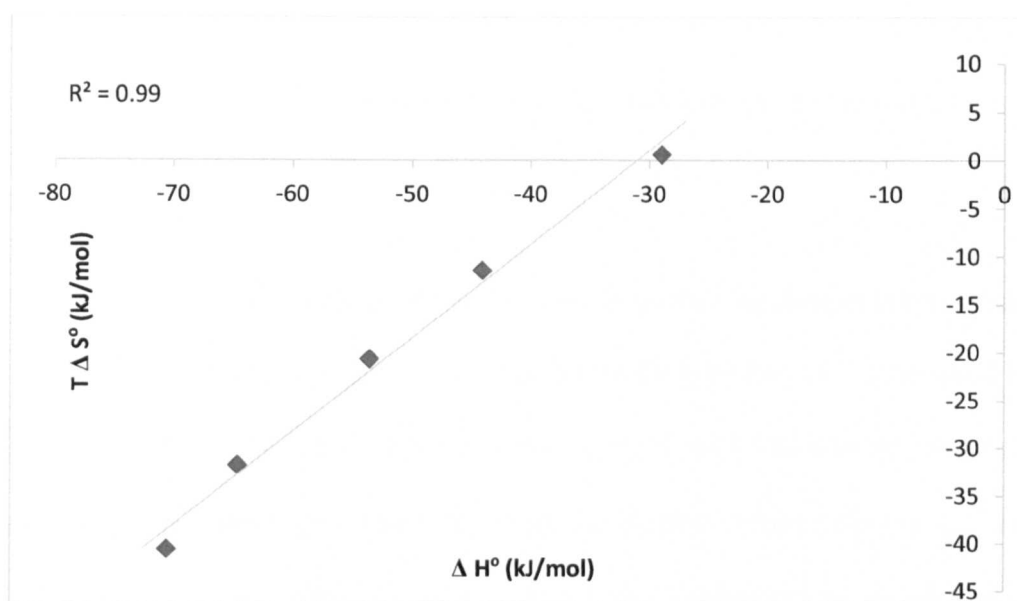


Figure 6.8 Apparent entropy-enthalpy compensation seen across the γ -lactones series when ΔH° is plotted against $T\Delta S^\circ$. A strong correlation is seen.

Seemingly, as chain length increases the interactions become less strongly enthalpy driven, with a less negative entropy change, whilst the free energy remains essentially constant. The term entropy-enthalpy compensation (EEC) is often applied when a linear relationship is shown between ΔH° and ΔS° .

Frequently what is described is an increasingly negative enthalpy change, at the cost of a loss of entropy (Krishnamurthy et al., 2006).

The concept of EEC has been questioned, with some dismissing the phenomenon as a statistical artefact (Sharp, 2001; Cornish-Bowden, 2002). It is true that for a number of techniques ΔS° and ΔH° are derived from the same temperature data series in a van't Hoff analysis and so the likelihood of such relationships being seen is high (Cornish-Bowden, 2002). However, by calorimetry enthalpy is directly measured, and K is derived, from which ΔG° and ΔS° are calculated.

Sharp (2001) looked at a range of experiments; including one dataset derived from calorimetric data and came to the conclusion that the apparent EEC seen could be “better explained by other causes”. Sharp showed that introducing randomly generated ΔH° values gave similarly strong linear relationships between ΔS° and ΔH° . Sharp suggested that the slope of the line represented a compensation temperature;

$$T_c = \frac{\delta \Delta H^\circ}{\delta \Delta S^\circ} \quad (\text{Equation 6.1})$$

A confidence interval in T_c could be derived from the estimated standard error of T_c from the fit of the line. If the experimental temperature was outside this confidence interval the correlation could be considered significant. By this method only one dataset was shown to have an “extra-thermodynamic” T_c value. The significance of this dataset was dismissed as it was formed from a linear

series of ligands and the ΔH° vs. ΔS° plot was considered to be the result of additivity.

The γ -lactone series could also be considered to come under this category and the EEC seen could be dismissed. Krishnamurthy et al. (2006), however, also studied a series of related ligands binding to bovine carbonic anhydrase and found a similar result to that found for this series of γ -lactones. As the chain length of the ligand was increased the change in enthalpy became less negative, whilst entropy increased. Their explanation for this was that initially the ligand chain atoms were tightly bound but, as the chain length was increased the first contacts were somehow destabilized, leading to an increase in the flexibility of the ligand within the binding interface and hence an increase in entropy was seen. To test this hypothesis for the γ -lactones, chemical shift perturbations were examined to see if they might reveal the subtle differences in binding of the γ -lactones to OBP3. Upon γ -lactone binding the shifts were subtle, indicating localised changes in the protein.

6.2.3 Binding of OBP3 to the γ -lactones by NMR

A $^1\text{H} / ^{15}\text{N}$ HSQC-TROSY of a 0.5 mM sample of OBP3 saturated with each of the five γ -lactones (at a concentration of 1 mM) was recorded (figure 6.9). The spectra are overlaid such that the effect of increasing chain length at each stage can be seen. The movement of a peak in the $^1\text{H} / ^{15}\text{N}$ HSQC-TROSY indicates the amide proton is in a different environment. This may be caused by being directly in contact with the ligand, or part of the ligand. NH chemical shift perturbations are also caused by structural changes to the protein, either in a localised part of the protein or across the entire protein, such as when ligand binding causes the protein to adopt a new conformation. A conformational change across the entire protein would lead to a large number of amide proton peaks moving, with some of them moving a great deal.

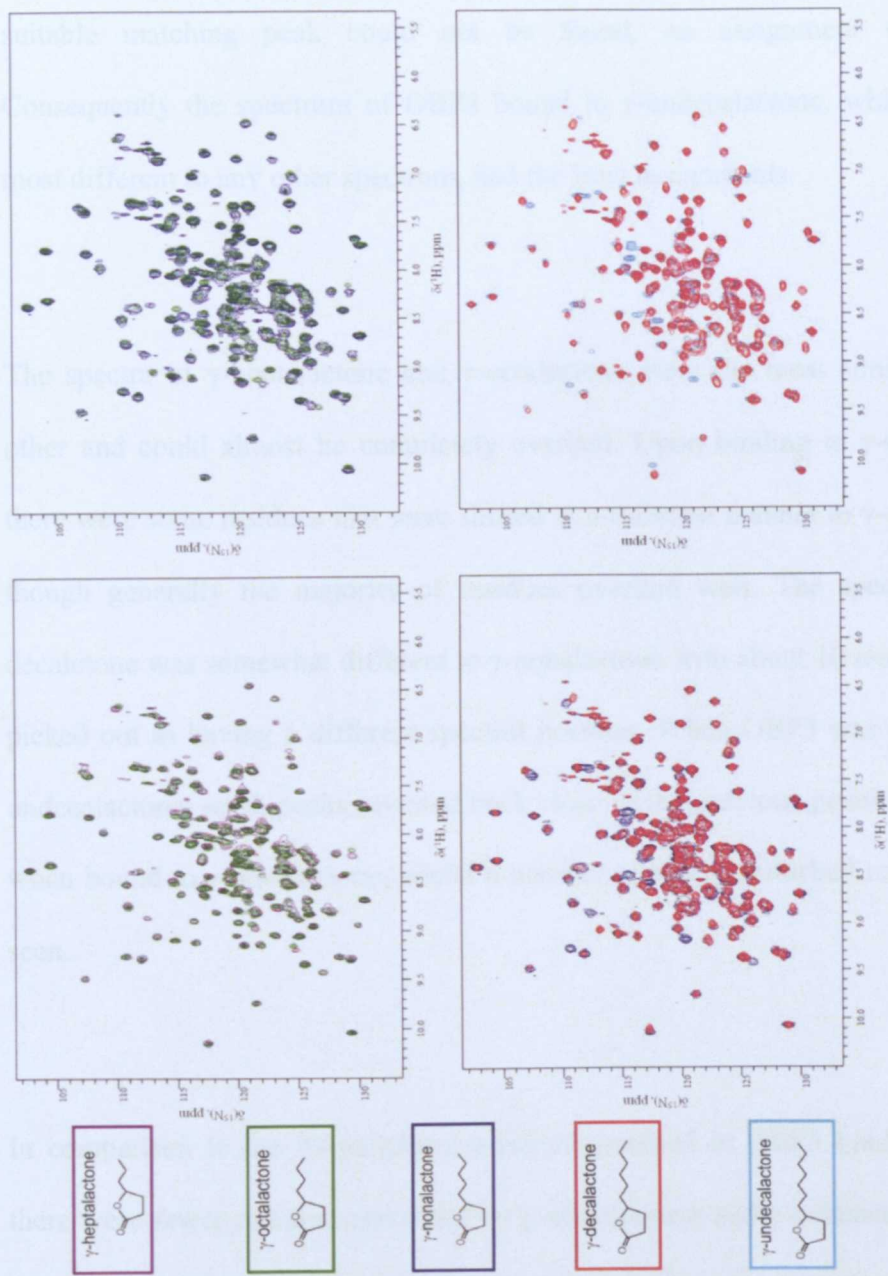


Figure 6.9 ^1H / ^{15}N HSQC-TROSY spectra of OBP3 bound to the γ -lactones. Top left: γ -heptalactone overlaid with γ -octalactone. Top right: γ -octalactone overlaid with γ -nonalactone. Bottom left: γ -nonalactone overlaid with γ -decalactone. Bottom right: γ -decalactone overlaid with γ -undecalactone.

Due to the number of spectra and the time that would be required, these spectra were not completely reassigned. The spectra were similar enough to unbound OBP3 and OBP3 bound to 2-isobutylthiazole (IBT) that a “nearest peak” assignment approach was carried out. Where assignments were ambiguous, or a suitable matching peak could not be found, no assignment was made. Consequently the spectrum of OBP3 bound to γ -undecalactone, which was the most different to any other spectrum, had the least assignments.

The spectra of γ -heptalactone and γ -octalactone were the most similar to each other and could almost be completely overlaid. Upon binding to γ -nonalactone there were some residues that were shifted in a different manner to γ -octalactone, though generally the majority of residues overlaid well. The spectrum of γ -decalactone was somewhat different to γ -nonalactone with about 10 residues easily picked out as having a different spectral position. When OBP3 was bound to γ -undecalactone, some peaks reverted back close to the previous position they held when bound to γ -nonalactone, whilst a number of newly perturbed residues were seen.

In comparison to the 39 perturbed residues involved in OBP3 binding to IBT, there were fewer residues perturbed by γ -heptalactone and γ -octalactone binding (30 and 29 respectively). γ -nonalactone and γ -decalactone perturbed 36 and 37 residues respectively on binding whilst γ -undecalactone had perturbed 40 residues, and a number of peaks were unassigned because the shift was too great to be followed. IBT has almost the same molecular weight as γ -octalactone.

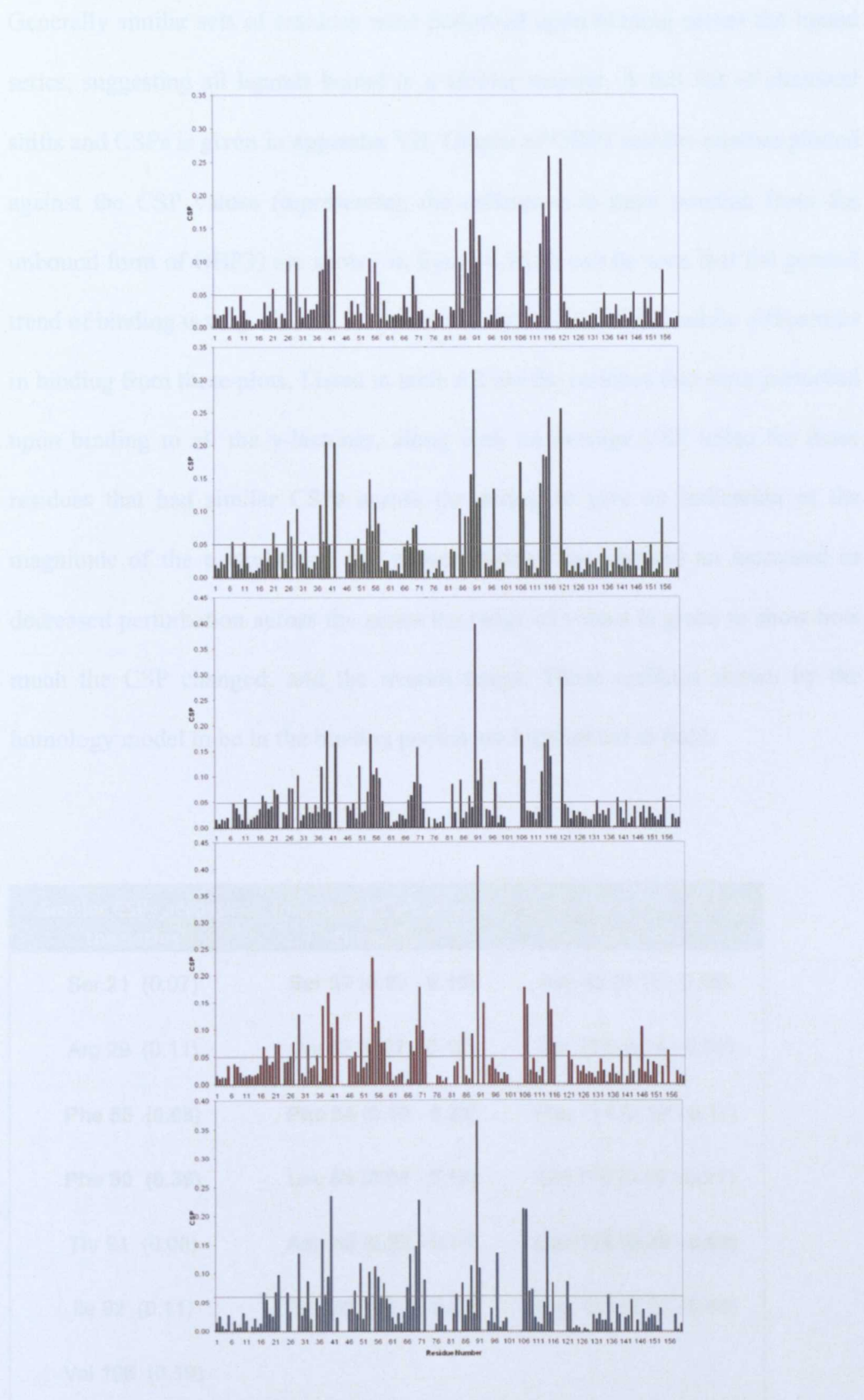


Figure 6.10 Chemical shift perturbations of OBP3 upon binding to each of the γ -lactones, from γ -heptalactone (top) to γ -undecalactone (bottom).

Generally similar sets of residues were perturbed upon binding across the ligand series, suggesting all ligands bound in a similar manner. A full list of chemical shifts and CSPs is given in appendix VII. Graphs of OBP3 residue number plotted against the CSP values (representing the difference in peak position from the unbound form of OBP3) are shown in figure 6.10. It can be seen that the general trend of binding is very similar; however it is difficult to extract subtle differences in binding from these plots. Listed in table 6.2 are the residues that were perturbed upon binding to all the γ -lactones, along with an average CSP value for those residues that had similar CSPs across the series, to give an indication of the magnitude of the perturbation. For those residues that showed an increased or decreased perturbation across the series the range of values is given to show how much the CSP changed, and the overall range. Those residues shown by the homology model to be in the binding pocket are highlighted in bold.

Similar CSPs across the series	Increasing CSPs across the series	Decreasing CSPs across the series
Ser 21 (0.07)	Ser 37 (0.09 - 0.16)	Asp 85 (0.10 - 0.08)
Arg 29 (0.11)	Gly 53 (0.07 - 0.10)	Thr 113 (0.14 - 0.05)
Phe 56 (0.08)	Phe 54 (0.10 - 0.23)	Phe 114 (0.19 - 0.17)
Phe 90 (0.36)	Leu 69 (0.08 - 0.15)	Gln 115 (0.19 - 0.17)
Thr 91 (0.08)	Asn 88 (0.08 - 0.11)	Leu 116 (0.26 - 0.08)
Ile 92 (0.11)	Tyr 120 (0.25 - 0.29)	Asp 155 (0.09 - 0.06)
Val 106 (0.19)		

Table 6.2 Residues that were perturbed more than 0.05 ppm upon binding to all the γ -lactones.

CSPs are given in brackets (in ppm). In bold are the binding pocket residues.

The most striking chemical shift perturbation was seen for Phe 90, which was highly perturbed (to a similar extent) by binding to each of the γ -lactones. This residue was largely unaffected when IBT bound to OBP3 (CSP = 0.03) and was exchanged with D₂O within 72 hours, unlike many other residues in the predicted binding pocket, raising the question of whether it was in the ligand binding cavity. This suggests that the γ -lactones either bind in a different area of the pocket, or when they bind they modulate the binding pocket, somewhat differently to when IBT binds. Phenylalanine is hydrophobic and hence would be expected to be affected by a hydrophobic interaction with a hydrophobic ligand. However, the residue was not more greatly affected by the more hydrophobic γ -lactones (those with a longer hydrocarbon chain) and in fact IBT has an estimated log P value of 2.51 (Rabe, 2004) which is higher than all the γ -lactones except γ -decalactone and γ -undecalactone (shown in figure 6.2).

The residues that were perturbed by binding to all the γ -lactones are shown in figure 6.12 (the residues that showed increased shifts are shown in green and those that decreased are shown in red). This figure shows those residues in relation to the other binding pocket residues (shown in grey). The perturbation of Tyr 120 across the series was also quite high and showed a trend of increasing as the ligand size increased. All CSPs for Tyr 120 were a lot higher than that recorded for OBP3 binding to IBT (0.10 ppm). It is possible that the hydroxyl group of Tyr 120 is hydrogen bonded to the ketone group on the lactone ring, which would be different in nature to a hydrogen bond formed between the tyrosine hydroxyl group and the nitrogen of the thiazole and may account for the different magnitudes of perturbation seen. This does not explain why Tyr 120 is

more highly perturbed as the chain length increases. This may be due to the fact that tyrosine is quite a large residue in the pocket and consequently must move in order to accommodate the larger and larger ligands or it may simply be the result of the increasing hydrophobicity of the ligand increasing the hydrophobic interaction. This is also likely to be the case for Phe 54 which shows increasingly large perturbations across the series.

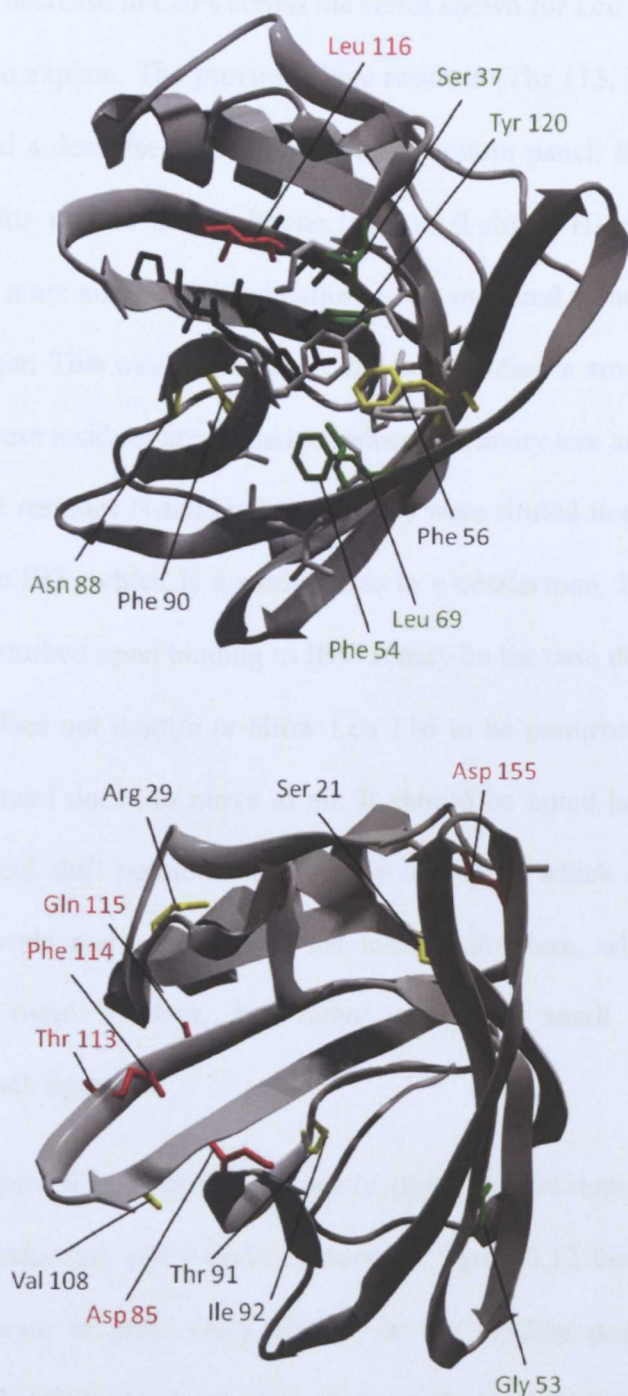


Figure 6.11 Residues that are perturbed more than 0.05 ppm upon binding to all of the γ -lactones. In the top panel the binding pocket residues are shown, looking from the “top” of the barrel. Residues in yellow have similar CSPs across the series. Those residues in green show increased CSPs as the chain length is increased whilst those in red show decreased CSPs as the chain length is increased. Below, using the same colour scheme are the residues that aren’t in the binding pocket but are perturbed upon binding to all the γ -lactones.

The significant decrease in CSPs across the series shown for Leu 116 is somewhat more difficult to explain. The previous three residues (Thr 113, Phe 114 and Gln 115) all showed a decrease in the CSP values (bottom panel, figure 6.11). This suggests that this part of the backbone (part of β -strand H) moves back to a position that is more similar to its situation in the unbound form of OBP3 as the ligand gets larger. This model assumes that upon binding a small ligand like γ -heptalactone, these residues are shifted to reduce the cavity size around the ligand. Whilst the three residues N-terminal to Leu 116 were shifted in a similar manner upon binding to IBT, which is a similar size to γ -octalactone, Leu 116 was not significantly perturbed upon binding to IBT. It may be the case that the shape and nature of IBT does not require or allow Leu 116 to be perturbed, such that this part of the β -strand does not move at all. It should be noted however, that the maximal chemical shift perturbations are only 0.26 ppm which are too small to indicate large scale rearrangement of the tertiary structure, which would also involve many more residues, but rather points to small adjustments to accommodate each ligand.

As noted in figure 6.9 a number of newly perturbed residues are seen upon binding to γ -decalactone and γ -undecalactone. In figure 6.12 the binding pocket residues are shown in grey. Only Phe 67 in the binding pocket was newly perturbed upon binding to γ -decalactone and γ -undecalactone (residues shown in claret). The perturbations seen (0.08 ppm and 0.13 ppm respectively) may suggest the ligands are in quite close proximity to this residue. The rest of the perturbed residues surround the pocket and the residues that are only perturbed upon binding to γ -undecalactone (shown in orange) are generally even more distal from the binding pocket. Most of the perturbations are very small; however the

perturbation for Glu 66 is 0.09 ppm, which is unsurprising as it is in close proximity to Phe 67. Ile 32 and Asp 72 have CSPs of 0.09 ppm upon binding to γ -undecalactone; these residues are 22 Å apart, on opposite sides of the pocket, suggesting that the protein is more open or stretched when accommodating these larger ligands. It may be the case that, as increasingly larger ligands are bound, the protein becomes more flexible in order to accommodate them. Previous studies have shown that MUP I becomes more mobile upon binding to 2-sec-butyl-4,5-dihydrothiazole (Macek et al., 2006), therefore the mobility of OBP3 was investigated using ^{15}N -Heteronuclear NOE experiments.

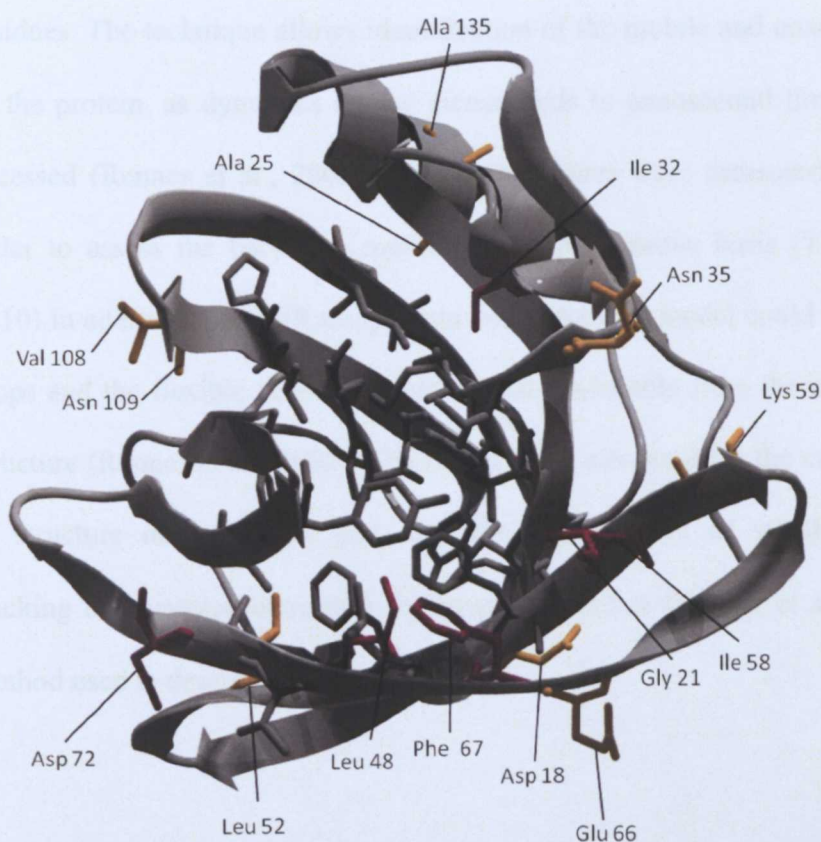


Figure 6.12 Residues perturbed upon binding to γ -decalactone and γ -undecalactone (claret) and residues only perturbed upon binding to γ -undecalactone (orange). The binding pocket residues are shown in stick form, only Phe 67 in the binding pocket is perturbed (shown in claret).

6.2.4 Heteronuclear NOE experiments to probe the mobility of OBP3

Protein mobility and flexibility are important in ligand recognition. It is possible to look at the dynamics of the protein backbone using heteronuclear NMR experiments (Clore et al. 1990). In the $^1\text{H} / ^{15}\text{N}$ Heteronuclear Nuclear Overhauser Effect (Het-NOE) experiment magnetization is transferred from the amide proton to the heteronucleus (^{15}N). The experiment measures the cross relaxation rate. The transfer rate depends on the dynamics and therefore the Het-NOE reports the motion of individual N-H bond vectors. Motions faster than the overall tumbling of the protein molecule show a decreased NOE intensity relative to the rest of the residues. The technique allows identification of the mobile and unstructured parts of the protein, as dynamics on the picoseconds to nanosecond timescale can be accessed (Renner et al., 2002). . Het-NOE values were measured for OBP3 in order to assess the backbone mobility on a per residue basis (Treweek et al., 2010). In addition the NMR assignment and homology model could be verified, as loops and the flexible termini should be distinguishable from the core secondary structure (Renner et al., 2002). The technique is also used for the cross-validation of structure determination and also for identification of suitable candidates (lacking unstructured elements) for structural studies (Renner et al., 2002). The method used is described in section 2.6.3.7.

The het-NOE values for OBP3 without a ligand bound are shown in figure 6.13. The errors bars are standard deviations calculated by the CCPNMR software programme and reflect the quality of the peak picked, as the reference and saturated peak are compared. For some weak peaks (such as Val 139) the standard

deviation values were greater than 1, these peaks were therefore excluded from the analysis.

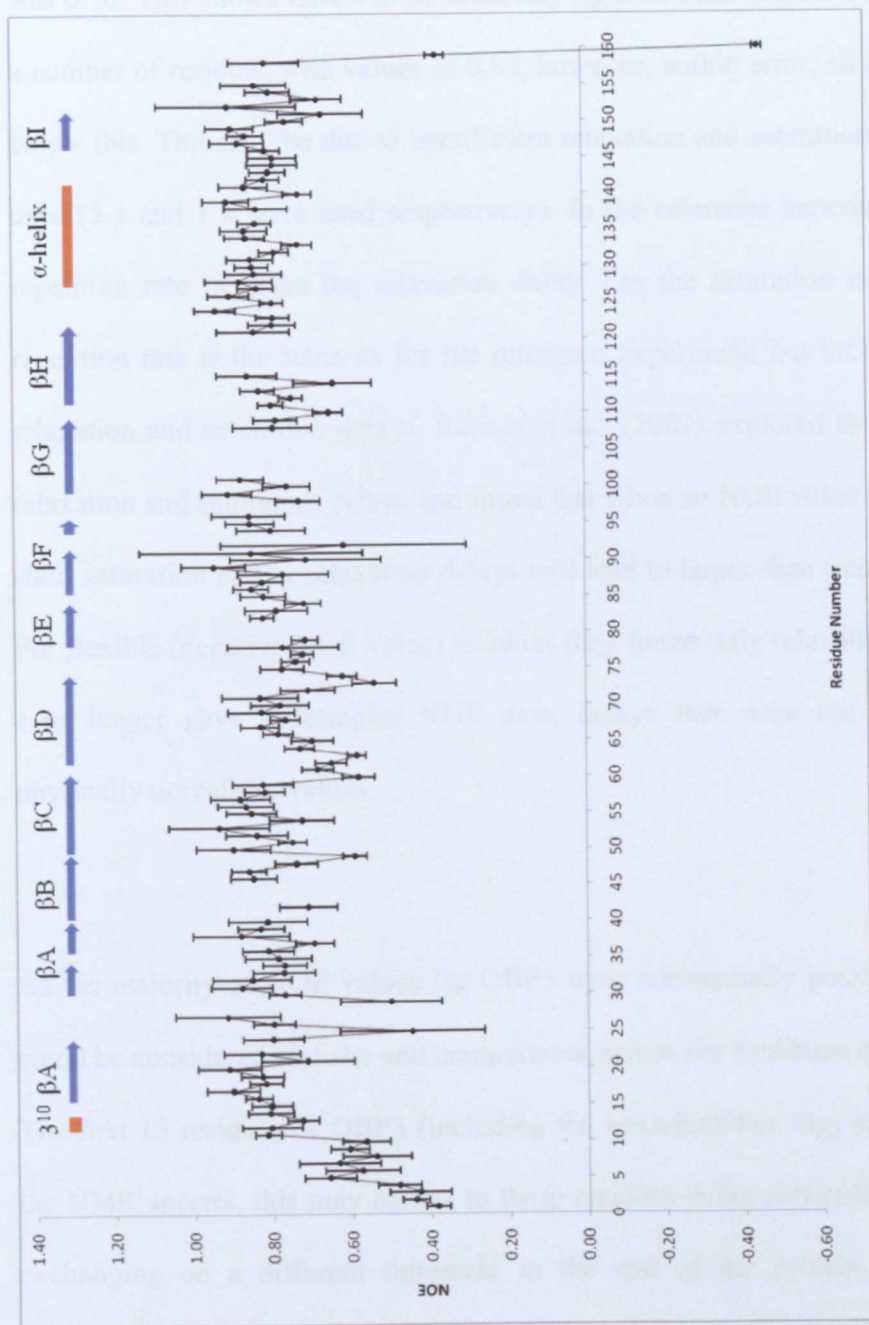


Figure 6.13 $^1\text{H} / ^{15}\text{N}$ het-NOE values of the residues of unbound OBP3. The secondary structure elements of OBP3 are shown along the top, with β -strands in blue and helices in orange. The more positive the value the more rigid the corresponding residue is.

Het-NOE values can go from negative values (indicating large mobility) to positive values (indicating rigidity). The theoretical maximum het-NOE value is quoted as 0.82 (Grasberger et al., 1993). The average het-NOE value for OBP3 was 0.76. This shows OBP3 to be relatively rigid overall. Within OBP3 there are a number of residues with values of 0.82, however, within error, all the values are below this. This may be due to insufficient relaxation and saturation delays being used (5 s and 1 s were used respectively). In the reference experiment the scan repetition rate includes the relaxation delay. For the saturation experiment the repetition rate is the same as for the reference experiment but includes both the relaxation and saturation delays. Renner et al. (2002) explored the effect of the relaxation and saturation delays and found that when an NOE value is positive too short saturation and/or relaxation delays will lead to larger than true NOE values. For flexible (negative NOE value) residues they found only relaxation delays of 5 s or longer gave meaningful NOE data, delays that were too short yielded physically unrealistic values.

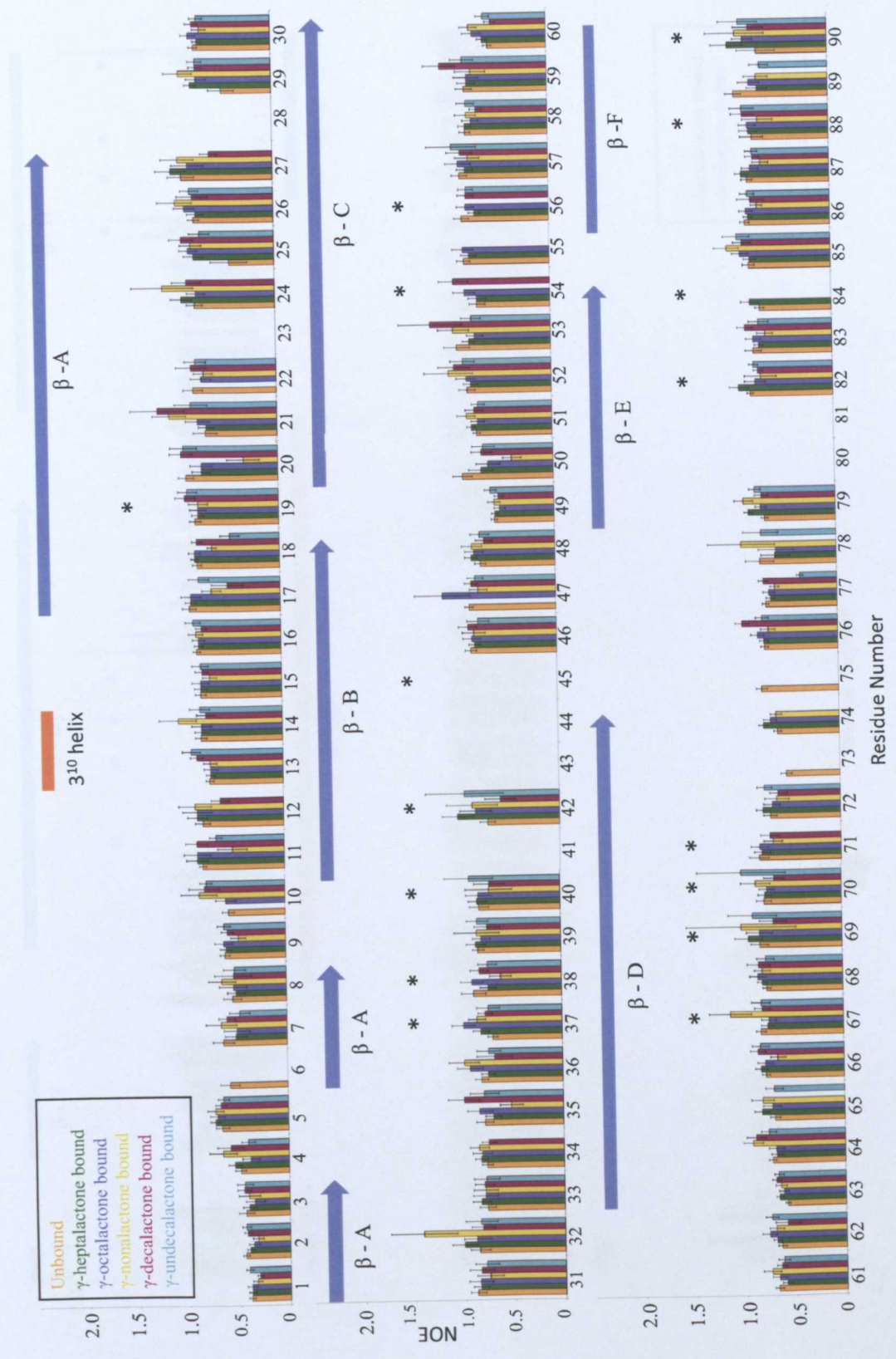
As the majority of NOE values for OBP3 were substantially positive the values could be considered realistic and comparisons across the backbone could be made. The first 13 residues of OBP3 (including the hexa-histidine tag) are not seen on the NMR spectra, this may be due to these residues being particularly flexible or exchanging on a different timescale to the rest of the protein. These initial residues are followed by residues 1-10 which had NOE values between 0.39 and 0.66, which is below the average value for OBP3, showing this part of the protein to be less rigid. This fits with the lack of secondary structure shown by the homology model. Some other residues also had lower values, indicating they are

not as rigid, particularly, Ala 25 (0.45), Arg 29 (0.50) and Glu 49 (0.59) which aren't part of the defined secondary structure elements.

Interestingly Lys 73 and Thr 74 were also less rigid than average, despite being modelled as at the end of a β -strand, possibly indicating they are not actually part of the β -strand. Thr 74 rapidly exchanged in the D₂O experiments (section 6.2.1); however, the other NH groups that had exchanged before the first experiment had been run do not appear particularly flexible in comparison with the rest of the protein. This suggests the rapid exchange is merely due to these being the most solvent accessible residues. The C-terminal alanine (Ala 160) unsurprisingly is flexible, shown by a negative NOE value. The penultimate residue, Gln 159, also shows increased flexibility (0.37), although the rest of the C-terminus seems to be quite rigid.

In general the NOE values of OBP3 are of the same magnitude described for MUP I (Zidek et al., 1999). In this study the NOE values of MUP I when ligand bound were measured and it was concluded that, contrary to what is generally observed, the protein mobility increased upon binding to a ligand. The γ -lactone series is ideal for looking at trends of increasing or decreasing protein flexibility. OBP3 samples were saturated with each of the γ -lactones and a ¹H-¹⁵N heteronuclear NOE experiment run for each, using the same parameters as for unbound OBP3. Figure 6.14 shows all the values plotted against residue number. From this figure it is clear that a dramatic increase or decrease in flexibility wasn't seen upon ligand binding, with values being within error of each other for most

residues. There were however some notable changes in mobility; Val 12, which is immediately before the 3_{10} helix, showed a decreased NOE value (~ 0.2 lower) when bound to γ -decalactone (with respect to all the other experiments (the residue was unassigned for the γ -undecalactone spectra), indicating that this residue was more flexible. This was also the case for Asp 18 which appeared to be more flexible when OBP3 was bound to γ -undecalactone. This residue is at the beginning of the first (N-terminal) β -strand. Ala 25 and Arg 29, which were noted as being more flexible than the rest of the unbound form of OBP3, showed increased NOE values (all of similar magnitude) when OBP3 was bound to the γ -lactone. Arg 29 is part of a region without regular secondary structure, however it appears upon binding that in this region only this residue becomes more rigid. Upon binding to γ -nonalactone Phe 67 (which is in the binding pocket) was shown to become particularly more rigid, although when bound to γ -decalactone and γ -undecalactone this increased rigidity was not maintained.



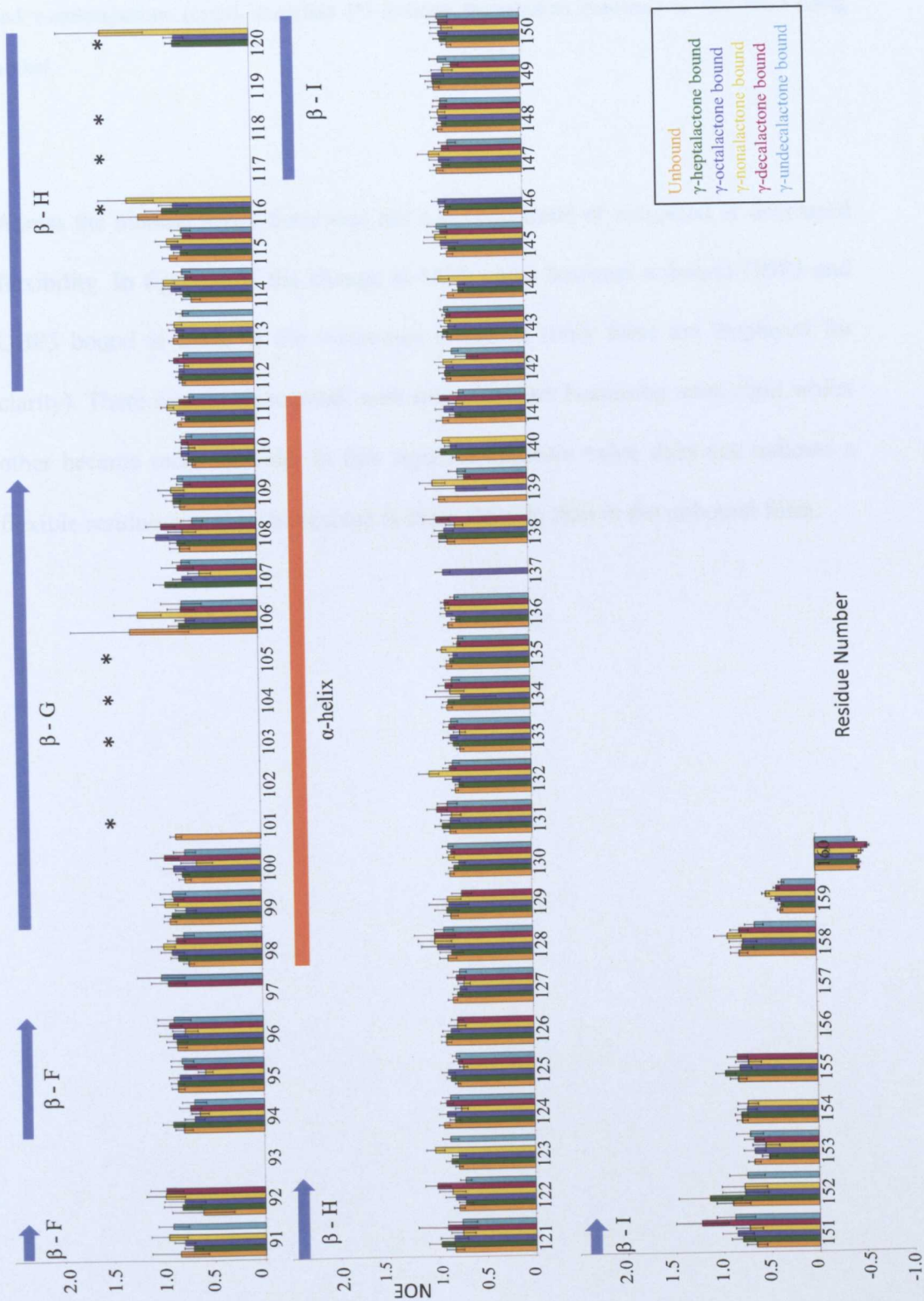
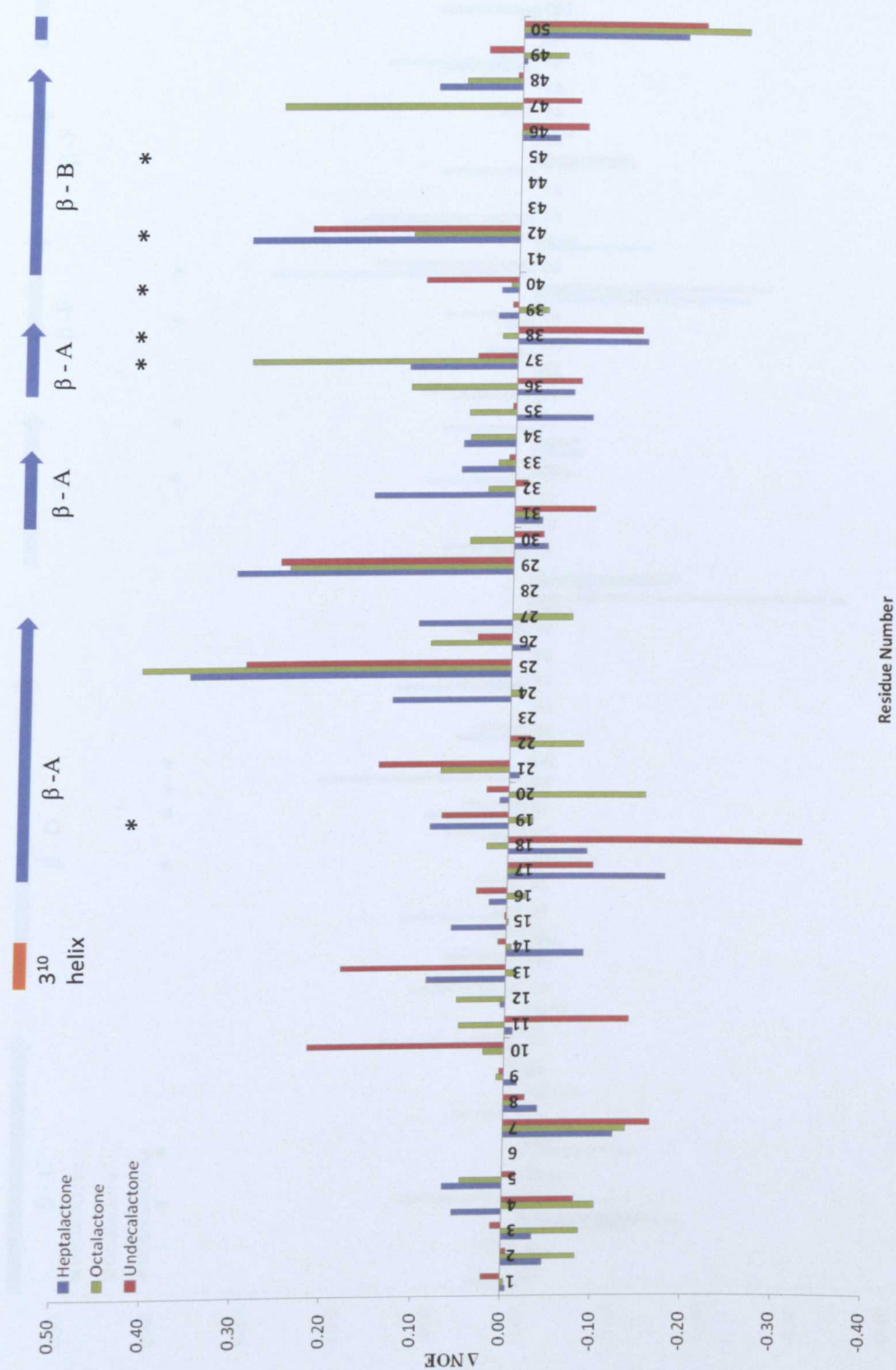
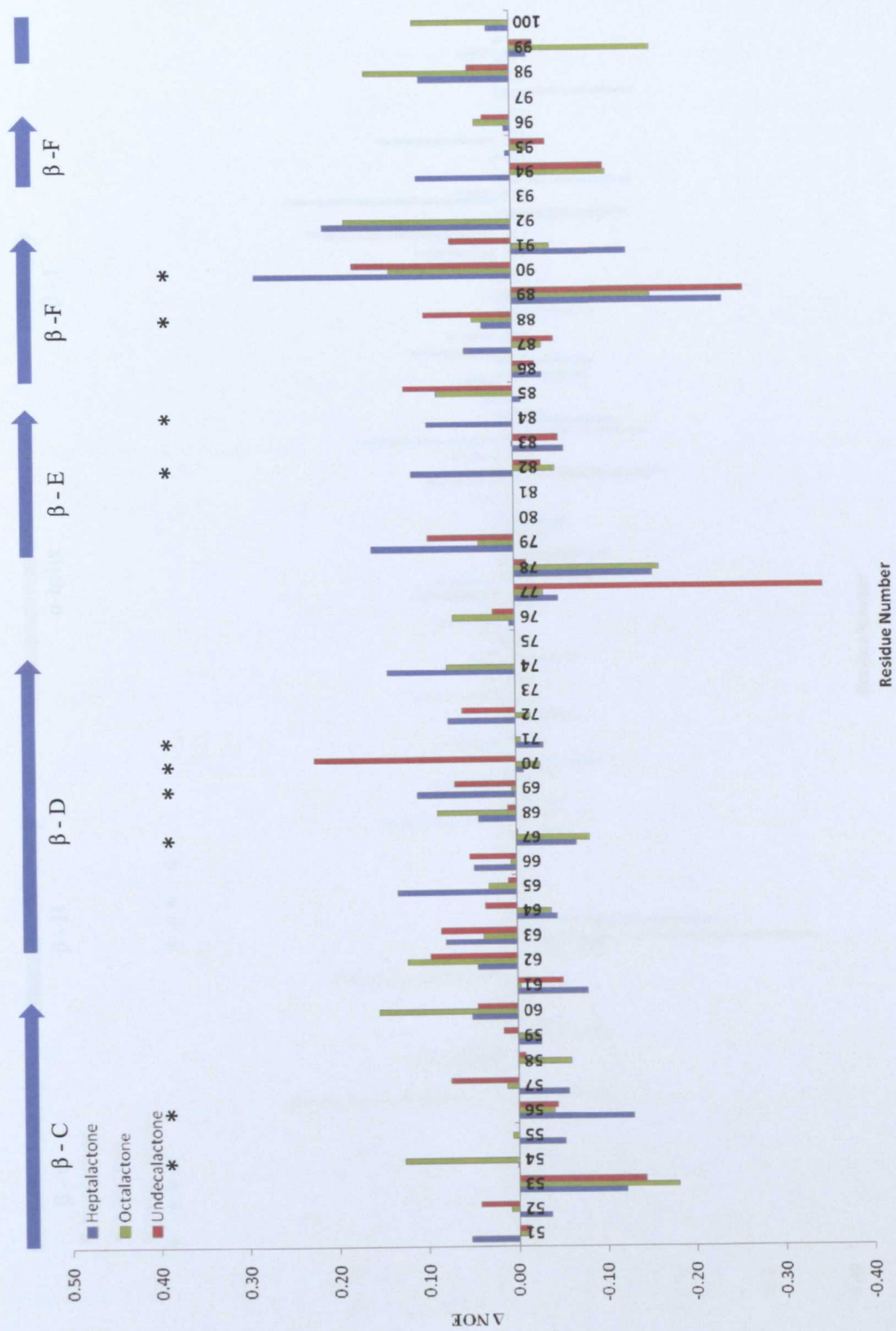


Figure 6.14 ^1H - ^{15}N het-NOE values for all residues of OBP3 when unbound (orange) and bound to; γ -heptalactone (green), γ -octalactone (blue), γ -nonalactone (yellow), γ -decalactone (maroon) and γ -undecalactone (cyan). Asterisks (*) indicate the residues predicted to line the binding pocket.

Across the binding series there was not a general trend of increased or decreased flexibility. In figure 6.15 the change in NOE value between unbound OBP3 and OBP3 bound to three of the γ -lactones is shown (only three are displayed for clarity). There was no clear trend, with some residues becoming more rigid whilst other became more flexible. In this figure a negative value does not indicate a flexible residue, just that the residue is more flexible than in the unbound form.





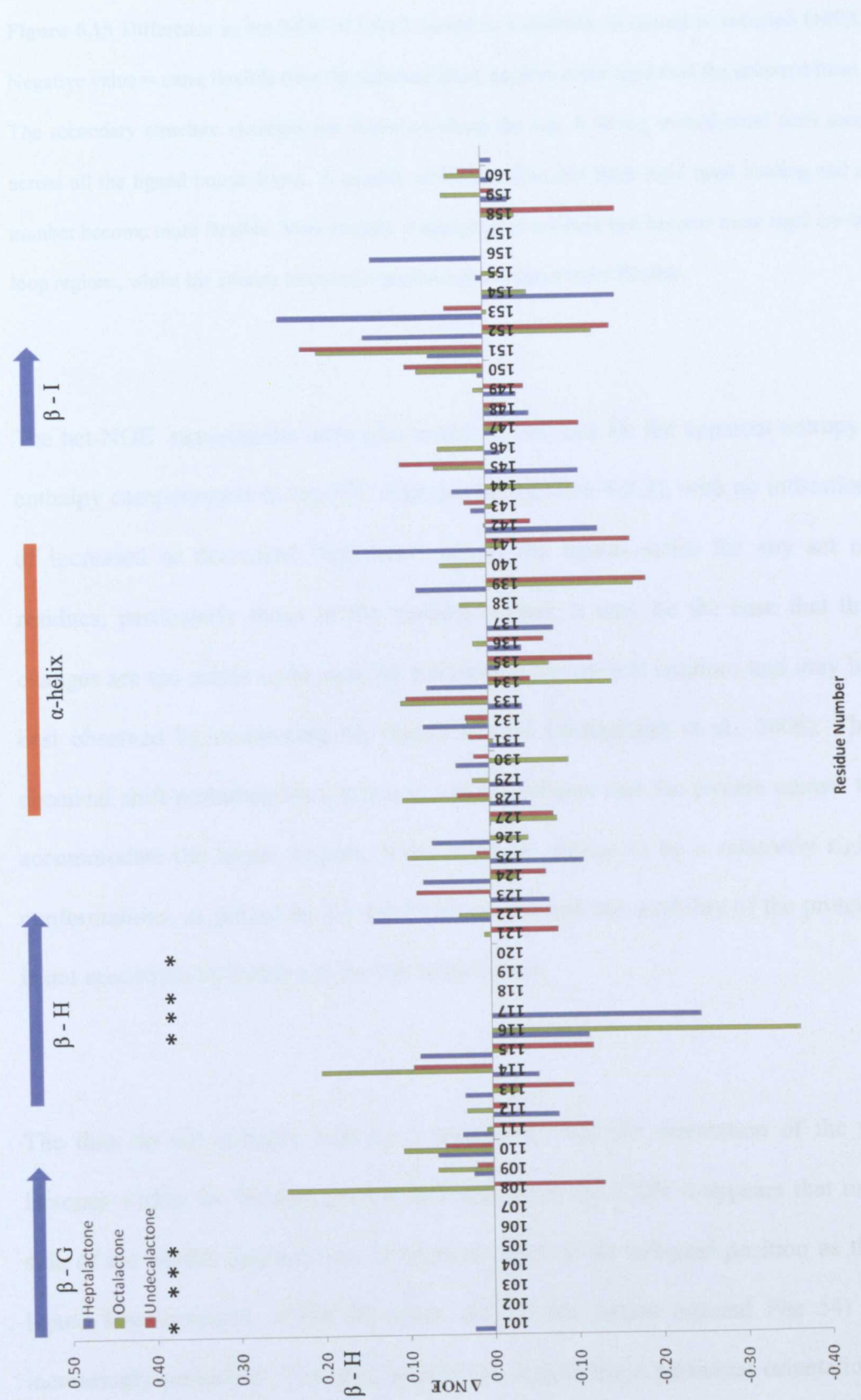


Figure 6.15 Difference in het-NOE of OBP3 bound to γ -lactones compared to unbound OBP3. Negative value = more flexible than the unbound form, positive more rigid than the unbound form. The secondary structure elements are displayed along the top. A strong overall trend isn't seen across all the ligand bound forms. A number of residues become more rigid upon binding and a number become more flexible. Very broadly it appears that residues that become more rigid are in loop regions, whilst the strands have more residues that become more flexible.

The het-NOE experiments were also unable to account for the apparent entropy-enthalpy compensation in the ITC experiments (section 6.2.3), with no indication of increased or decreased “tightness” across the ligand series for any set of residues, particularly those in the binding pocket. It may be the case that the changes are too subtle to be seen by monitoring the protein residues and may be best observed by monitoring the ligand instead (Stöckmann et al., 2008). The chemical shift perturbations (section 6.2.4) do indicate that the protein adjusts to accommodate the larger ligands, however these appear to be a relatively rigid conformations, as judged by the het-NOE values, and any mobility of the protein is not accessible by looking at the het-NOE values.

The data do not strongly indicate a particularly specific orientation of the γ -lactones within the binding pocket, although from the CSPs it appears that one side of the pocket (around Leu 116) moves back to an unbound position as the ligand size increases, whilst the other side of the pocket (around Phe 54) is increasingly perturbed. This may suggest the ligand has a particular orientation, rather than rotating freely, and may possibly hydrogen bond to Tyr 120, which would limit the likely positions that the ligand may occupy in the binding pocket.

To explore this further the binding of a set of three acetate ester isomers was observed by ITC to see what impact ligand shape and possibly orientation in the binding pocket had on the thermodynamic parameters and affinities and explore if interactions with specific residues were likely to be taking place.

6.2.5 Binding of the acetate esters by ITC

The acetate esters used were isobutyl acetate (IBA), sec-butyl acetate (SBA) and tert-butyl acetate (TBA). They are relatively small, having a molecular weight of 116.2, which is less than γ -heptalactone. They all possess a sweet smelling odour, often described as fruity (Howard, 1993). IBA occurs naturally at low concentrations in raspberries and pears, whilst TBA is described as having a blueberry like odour.

The acetate esters have previously been used by this laboratory in experiments studying their binding to OBP1 using Atmospheric-pressure chemical ionization mass spectrometry (APCI-MS) based techniques, as they proved to be most useful due to their rapid transit through the device (Borysik et al., 2010). All the acetate esters used showed good binding, by ITC, to OBP1. Dissociation constants recorded ranged from 0.64 μ M for TBA to 1.06 μ M for IBA. The thermodynamic parameters for each reaction were similar and indicated enthalpy driven interactions were taking place.

For OBP3 the ITC experiments were carried out in the same manner described previously (section 5.2.2). The resulting binding curves are shown in figure 6.16. Somewhat surprisingly, no interaction with TBA was seen. The experiment was repeated using a higher concentration of the odour but still no binding occurred.

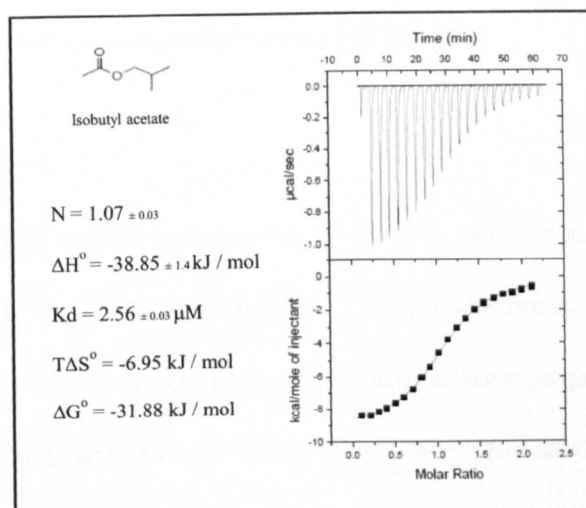
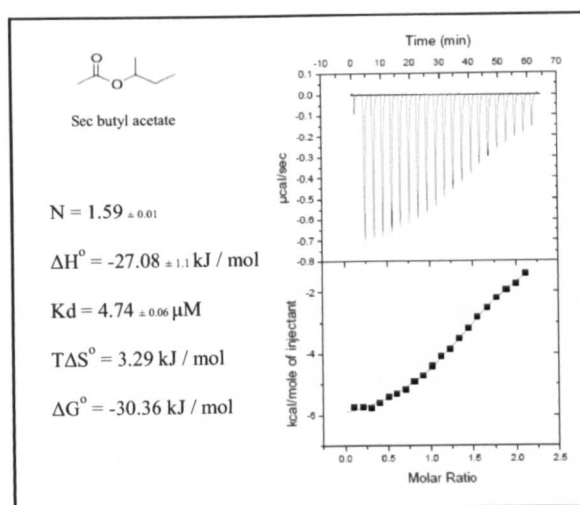
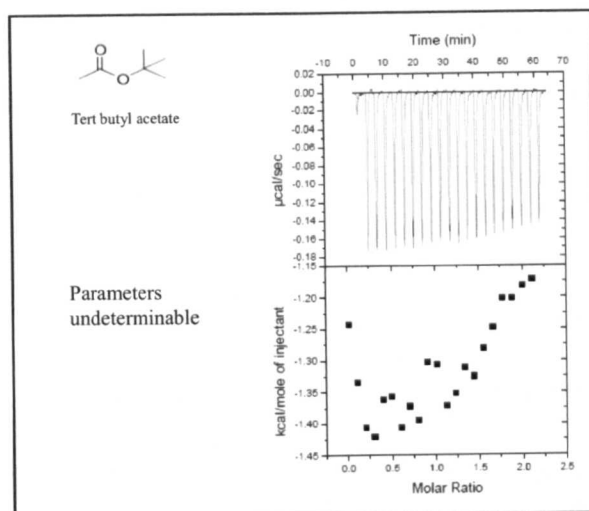


Figure 6.16 ITC of OBP3 binding to the acetate esters; Tert butyl acetate (TBA), Sec butyl acetate (SBA) and Isobutyl acetate (IBA).

The configuration of TBA is the most different to the other acetate esters and it is possible that some steric hindrance was taking place. This is curious however as TBA showed such strong binding to OBP1. The ITC experiments on OBP1 were performed at 21 °C, rather than the 25 °C used for OBP3, so direct comparisons cannot be made. However, qualitatively it does suggest different specificity of binding between these subtypes is occurring. One possibility is that the larger binding pocket of OBP3 is unable to contain TBA due to steric hindrance preventing the odour being “tethered” in the pocket. TBA is also much more volatile than the other acetate esters which may contribute to the inability of OBP3 to retain this odour.

SBA showed binding to OBP3 with reasonable affinity ($K_d = 4.74 \mu\text{M}$), however the reaction did not reach saturation. Overall the interaction was favourable with a negative ΔG° value. The enthalpy term was also large and negative, which suggests an enthalpy driven reaction, however a slight increase in entropy was measured, suggesting an entropic component to binding.

The binding of IBA to OBP3 was more favourable, and strong 1:1 binding was seen. A reasonably high binding affinity of $2.56 \mu\text{M}$ was measured. This is almost double the affinity measured for SBA. This reaction has a large negative enthalpy value and a small negative entropy value, which shows the reaction to be enthalpy driven.

6.3 Conclusions

In this chapter the solvent accessibility of free OBP3 was explored using D₂O exchange experiments. Unsurprisingly a number of glycine residues in loop regions were exchanged rapidly, whilst a set of core residues (particularly in β -strands C and D) remained unexchanged after 72 hours. A large number of unexchanged residues were in the binding pocket. Tyr 120 was a notable exception as it was amongst the residues exchanged within 10 minutes. D₂O exchange experiments with ligand bound OBP3 proved impossible due to the volatility of the odours so further analysis was not undertaken.

The γ -lactones were used to explore ligand binding to OBP3 in a systematic manner. As the hydrocarbon chain length was increased the size and hydrophobicity of the ligands increased. Across the series the free energy of binding remained consistent. This however was accompanied by somewhat counter intuitive entropy-enthalpy compensation. As ligand size is increased, one might expect to see an increased number of Van der Waals contacts and hence a more negative enthalpy. In fact the enthalpy values became less negative. This was accompanied by apparent compensation, with an increasingly negative ΔS° value. It seems the larger the γ -lactone the less stable the binding interaction was.

Chemical shift perturbations were measured for each of the γ -lactones. Interestingly a number of residues had increasing CSPs across the series, whilst

others showed decreasing CSPs. Looking more closely at the binding pocket residues it was noted that Phe 90 was perturbed by the γ -lactones but not by IBT. Additionally Ser 37, Phe 54, Leu 69, Asn 88 and Tyr 120 were increasingly perturbed with increasing lactone size. Leu 116 and other nearby residues in strand H were perturbed less by the larger lactones. This indicates that the OBP3 binding pocket makes adjustments to accommodate the larger ligands. Residues away from the binding pocket were found to be only perturbed by the largest γ -lactones looked at.

The $^1\text{H} / ^{15}\text{N}$ Het-NOE data showed OBP3 to be quite rigid, with only some residues in loop regions displaying any mobility. The N and C termini appeared to be more flexible than the remainder of the protein. Although binding to the γ -lactones had what appeared to be a pronounced effect on some residues (such as Arg 29) these isolated changes should not be over-interpreted. Large scale increased or decreased flexibility of stretches of residues was not seen.

Finally, the binding of three acetate esters were examined and it was highlighted that these isomers had differing binding affinities (in fact tert-butyl acetate didn't bind at all). Despite being the same size and having the same functional groups, the different shapes and spatial arrangements of the molecules lead to markedly different binding. OBP,1 which shows high affinity binding to the acetate esters overall, does not show such differences.

7. PROTEIN ENGINEERING OF THE OBP3 BINDING POCKET

7.1 Introduction

In chapters 5 and 6, the ligand binding of wild type OBP3 was explored by ITC and NMR. It was found that the residues in the binding pocket play an important role in binding odours. The size of the pocket is also significant, particularly when accommodating larger ligands like γ -undecalactone. To learn more about the OBP3 binding pocket and the contributions of individual residues and the size of the pocket, alterations were made and the resulting effects on binding were monitored by ITC. The mutants constructed fall into two sets which will be discussed in turn. Firstly the importance of the aromatic nature of the pocket was examined. This was performed by exchanging aromatic residues in OBP3 for alanine residues. These mutations also increased the size of the binding pocket. Secondly specific mutations were made as a result of comparing the predicted binding pocket of OBP3 and the binding pocket of OBP1 to see if OBP3 could be “tuned” such that it was able to more strongly (or weakly) bind to particular ligands.

7.1.1 The aromatic nature of the OBP3 binding pocket

The ligand binding pockets of vertebrate OBPs are highly hydrophobic in nature (Bianchet et al. 1996; Spinell et al. 1998; Paolini et al. 1998; White et al. 2009). The binding pocket of OBP3 is no exception and like other OBPs, it has a large number of aromatic residues, with 32% of the residues in the predicted ligand pocket being tryptophan (1), tyrosine (2) or phenylalanine (5). These three amino

from the homology model are shown in yellow. Identical residues are indicated by an asterisk (*), equivalent residues by a colon (:) and residues of the same nature by a dot (.).

All but two of the residues in the OBP1 pocket align with pocket residues in OBP3. It was decided to mutate the four conserved phenylalanine residues. In addition, the conserved tyrosine (Tyr 84) was also mutated, as well as the extra tyrosine residue (Tyr 120) found only in OBP3, and suggested from the data in previous chapters to play a possible role in hydrogen binding to ligands. The tryptophan residue was not mutated as this would have affected the ability to monitor the protein using UV absorbance (for example during protein purification) and it had not shown any significant involvement in binding based on the NMR data discussed in previous chapters (CSPs of 0.02 – 0.04 ppm). The intended sites for mutation are underlined and in bold in figure 7.1 and shown on the homology model of OBP3 in figure 7.2.

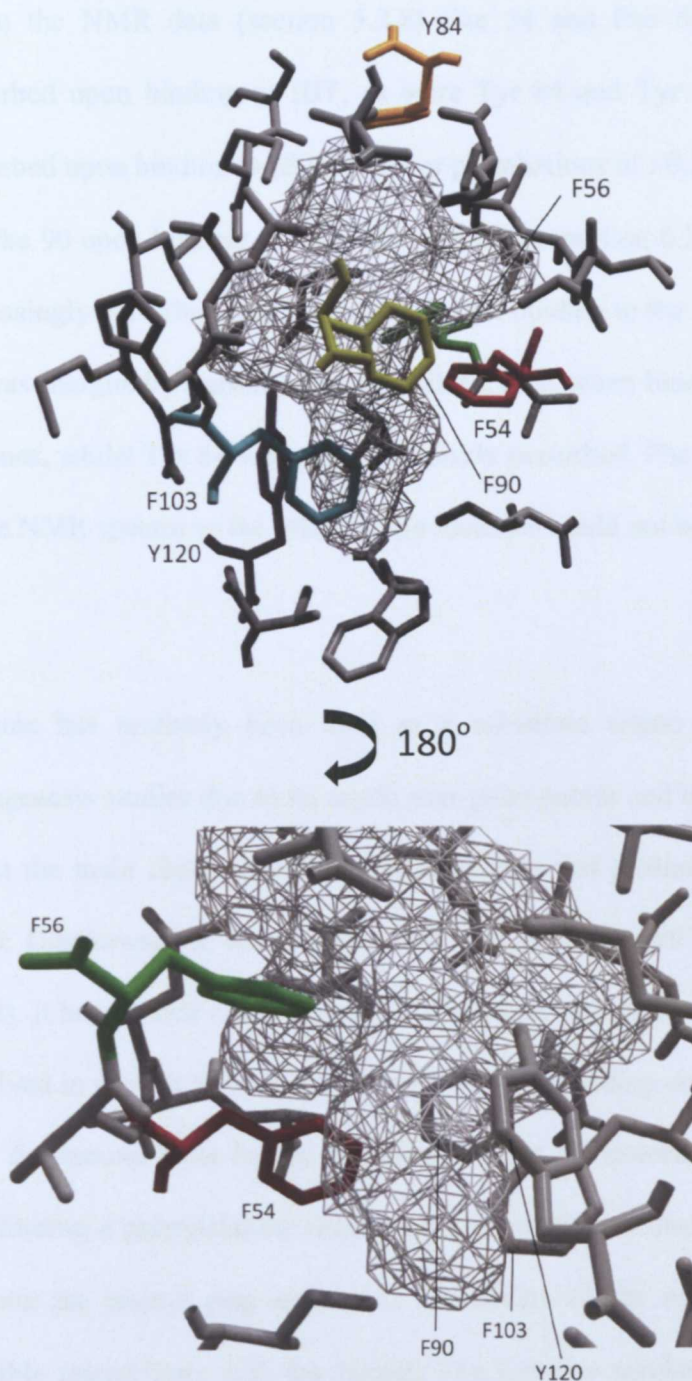


Figure 7.2 The binding pocket of OBP3 with the aromatic residues targeted for mutation highlighted. The cavity is denoted by the hatching. In the upper pane the pocket is orientated so that Tyr 84 is at the apparent opening of the pocket. The lower pane is rotated to show the other side of the pocket as indicated, giving a closer view of Phe 54 (red) and Phe 56 (green) within the binding pocket. Other nearby aromatic residues are also labelled but not coloured. It is clear that the aromatic residues selected are distributed around the pocket.

From the NMR data (section 5.2.8) Phe 54 and Phe 56 were shown to be perturbed upon binding to IBT, as were Tyr 84 and Tyr 120. Phe 90 was not perturbed upon binding to IBT; however perturbations of ~0.36 ppm were recorded for Phe 90 upon binding to all of the γ -lactones (section 6.2.4). Tyr 120 was also increasingly perturbed (0.25-0.29 ppm) upon binding to the longer γ -lactones. Phe 54 was marginally perturbed to a similar extent when binding to each of the γ -lactones, whilst Tyr 84 was not significantly perturbed. Phe 103 was not assigned in the NMR spectra so the affect of the mutation could not be anticipated.

Alanine has routinely been used as a substitute amino acid in site directed mutagenesis studies due to its small, non-polar nature and the fact that it does not affect the main chain conformation, as glycine and proline do (Ashkenazi et al. 1990; Dembowski & Kantrowitz 1994; L  fvre et al. 1997; Van Petegem et al. 2008). It has no side chain beyond the β -carbon, which means it is rarely *directly* involved in protein function (for example ligand binding or catalysis). In addition it is the second most highly abundant residue in proteins (Doolittle 1989). By substituting a phenylalanine residue with an alanine residue the major effect is to remove the phenyl ring and hence the ability of the residue to contribute to possible interactions with the ligand. The tyrosine residues are different to the phenylalanine residues as they have a hydroxyl function group and therefore can directly interact with the ligand. The substitution of a tyrosine with an alanine also removes the possibility of stacking interactions and additionally the possibility of the ligand being able to form hydrogen bonds. Both substitutions lead to a decrease in the side chain size, and hence an increase in the binding pocket size.

The other possible difference introduced is a change in the hydrophobicity of the residue. Phenylalanine, in particular, often plays a role in recognition of hydrophobic ligands due to its hydrophobic nature (Betts & Russell 2003). Ranking amino acids in terms of hydrophobicity is not straight forward. Various scales have been constructed; some based on the physiochemical properties of the side chains (Wolfenden et al. 1981; Kyte & Doolittle 1982) and some on the tendency for a particular amino acid residue to occur in the interior of proteins (Janin, 1979; Rose et al., 1985). For the purpose of these substitutions the physiochemical properties are most relevant. This places tyrosine as being less hydrophobic than alanine (Kyte & Doolittle 1982). Depending upon the scale used, phenylalanine is placed as both more and less hydrophobic than alanine. However, on balance, alanine was considered a reasonable amino acid to substitute for the aromatic amino acids.

7.2 Results and Discussion

7.2.1 Production of the aromatic mutants

All the aromatic mutants were expressed and purified in the same manner described for the wild type protein (sections 2.2 and 2.3). High levels of expression were achieved for all mutants and purification from the soluble fractions was possible with all the mutants, suggesting they were all likely to be stable. The overall yield of Y84A was somewhat lower than for the other mutants (around 30% lower) which indicated that this was the least stable mutant. The

aromatic mutants were prepared with the assistance of Miss Ina Tögel (technical intern).

The protein produced after purification of the mutants is shown in figure 7.3 and shows all the mutants to be approximately the same molecular weight, having bands between 16.5 kDa and 25 kDa. All bands appear to be clean with no signs of contamination from other proteins or degradation products.

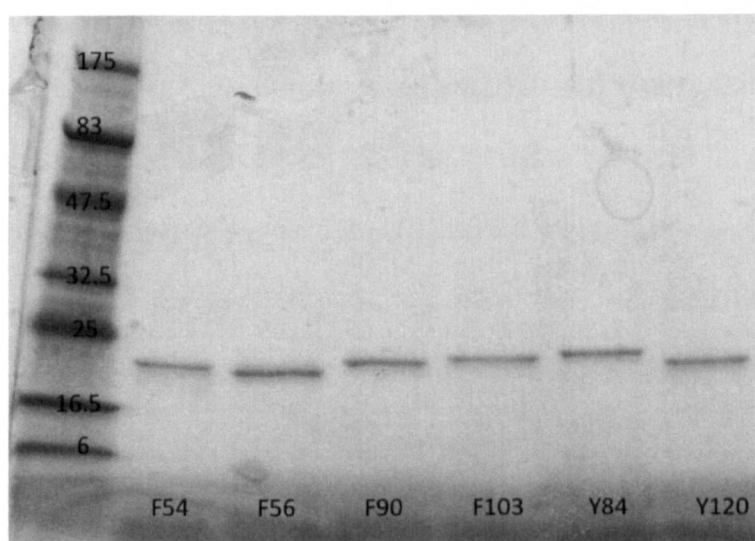


Figure 7.3 SDS-PAGE of the aromatic mutants demonstrating they were all free from contaminants and the expected molecular weight (~19.9 kDa)

To confirm the mutants were the expected molecular weight ESI-TOF mass spectrometry was carried out as described in section 3.5.1 and the molecular weights determined were compared to the wild type. All mutants matched the

theoretical change in molecular weight change from changing the original Phe or Tyr to an Ala confirming that the proteins had been correctly produced.

7.2.2 ITC of the aromatic mutants with 2-isobutylthiazole

To assess the functionality of the mutants, ITC was carried out as described previously (section 5.2.2). Initially binding to IBT was looked at. It was hoped that, as IBT was a ligand with high affinity to OBP3, even if a particular mutation made a large difference to the ability of OBP3 to bind the ligand it should be the case that enough binding to IBT would be retained to be able to derive binding parameters. The dissociation constants calculated for each of the OBP3 mutants is shown in the upper pane of figure 7.4. The inclusion of Y84A in the analysis makes it difficult to appreciate the values and more subtle differences in binding affinity recorded for the other mutants, therefore the lower pane of figure 7.4 shows the dissociation constants of the aromatic mutants, excluding Y84A.

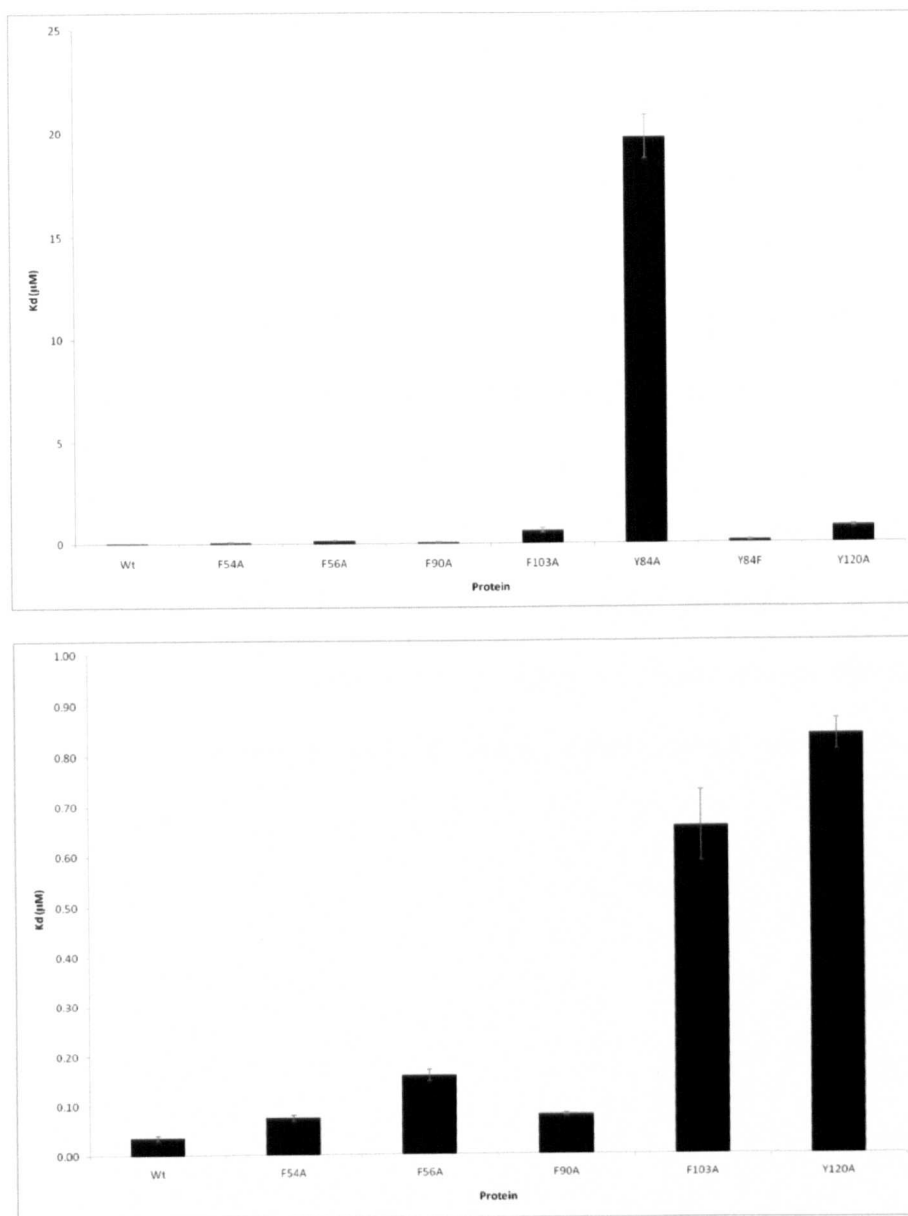


Figure 7.4 Upper pane: The dissociation constants of each of the aromatic mutants for IBT. The wild type value is included for comparison. Clearly Y84A binds much more weakly. Lower pane: The dissociation constants of each of the aromatic mutants for IBT excluding Y84A. The error bars represent the standard error of the mean of duplicate experiments.

It is most striking that Y84A has a dramatically reduced affinity for IBT. The binding of this mutant is completely different to that of the other mutants; the

binding curve produced was difficult to fit and lacked a sigmoidal shape. Figure 7.5 shows the original fitted binding curves for the aromatic mutants overlaid (along with the wild type curve) and demonstrates the weak binding shown by Y84A. All mutants had a lower stoichiometry value (n -value) than the wild type; this may be a result of an error in correcting the extinction coefficients of the mutants or perhaps due to more inactive protein being produced for the mutants than the wild type. Although affinity values are not within error of each other (with the exception of F54A and F90A) it should be noted that high affinity interactions are difficult to measure accurately, particularly at the protein concentration used (30 μ M), however figure 7.5 demonstrates that there are a number of data points on each slope which should allow more accurate measurement.

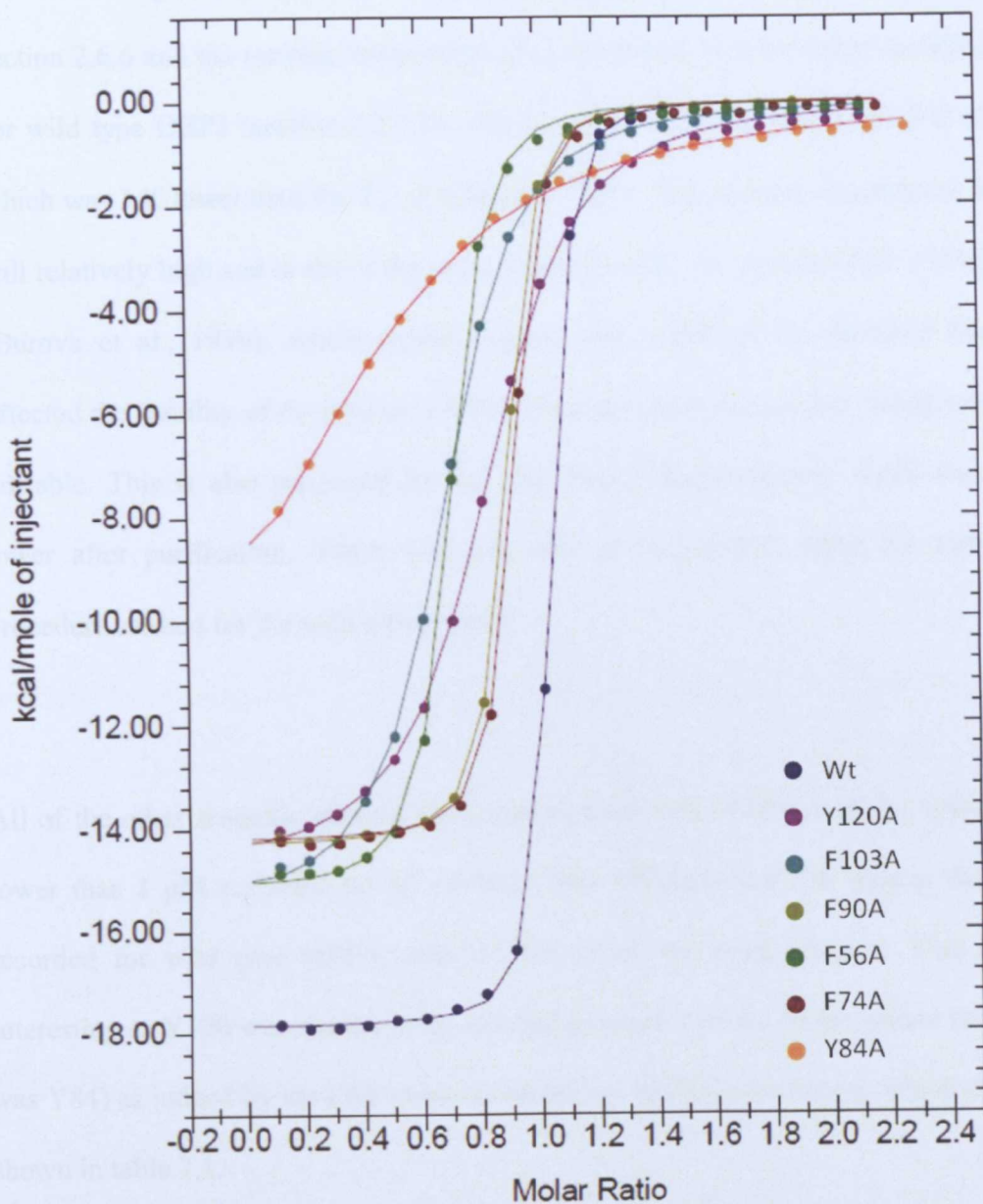


Figure 7.5 Fitted example ITC curves for each of the aromatic mutants binding to IBT. All show reduced affinity compared to the wild type however Y84A has a completely different, weaker, binding curve. The colours used correspond to the colours used on the model in figure 7.2.

The data indicated that Y84A warranted further investigation to determine if the mutation was particularly destabilising and therefore this was the cause of the

weaker binding. Differential calorimetry was therefore carried out as described in section 2.6.6 and the melting temperature (T_m) compared with the value recorded for wild type OBP3 (section 3.2.5.2). The T_m of Y84A was found to be 343 K which was 3 K lower than the T_m of wild type OBP3. This melting temperature is still relatively high and is above the value found by DSC for porcine OBP (342K) (Burova et al., 1999), which would suggest that, although the mutation has affected the stability of the protein a little, it has not made the protein completely unstable. This is also supported by the fact that, although protein yields were lower after purification, Y84A was still able to be purified using the same procedure as used for the wild type protein.

All of the other aromatic mutants continued to bind well to IBT, with K_d values lower than 1 μ M recorded for all of them. The affinities were all weaker than recorded for wild type OBP3, with Y120A being the most affected. This is interesting as Y120 was not the most affected aromatic residue in the pocket (nor was Y84) as judged by the CSP values recorded by NMR spectroscopy, which are shown in table 7.1.

Aromatic residue	Chemical shift Perturbation
Phe 54	0.19
Phe 56	0.10
Phe 90	0.03
Phe 103	-
Tyr 84	0.14
Tyr 120	0.10

Table 7.1 The chemical shift perturbations upon binding to IBT of the aromatic residues mutated in this chapter.

Tyr 120 however exchanged rapidly during the D₂O experiments (section 6.2.1), which suggested it was solvent accessible and possibly able to hydrogen bond to water molecules that were present in the pocket. Phe 54 was the aromatic residue with the greatest chemical shift change in the pocket upon binding to IBT, however F54A showed only slightly weakened binding. In fact the binding of F56A was weaker still despite this residue not being as perturbed upon binding to IBT. Figure 7.2 shows a view of these residues within the pocket, along with Thr 55 which connects them along the protein backbone. The phenyl rings of both Phe 54 and 56 point into the cavity. It may be the case that in the wild type protein Phe 54 takes part in a stronger hydrophobic interaction with IBT than Phe 56 does, due it being closer and more accessible to the ligand. In F54A the mutation makes the pocket larger and may allow Phe 56 to move into a position where it can take part in stronger interactions which compensate for the lack of the phenyl ring of Phe 54. Phe 103, Tyr 120 and Val 40 and 42 are also relatively closer to Phe 54 than to Phe 56 so may also be able to compensate to a certain degree.

When Phe 56 is mutated to an alanine, Phe 54 is unable to compensate for the mutation, perhaps due to the fact it already has a strong interaction with the ligand and therefore the additional small hydrophobic interaction Phe 56 was contributing, as shown by its CSP value, is simply lost.

Phe 90 is on the opposite side of the binding cavity to Phe 54 and 56, and showed no significant perturbation upon binding to IBT, it is therefore unsurprising that F90A showed little change in binding affinity compared to the wild type protein. Some difference was seen however, and this may reflect both the removal of the

contribution of Phe 90 to the hydrophobic nature of the pocket and possibly the increased size of the pocket. The likely compensation seen for F74A indicates the protein is able to accommodate the ligand well, despite the larger pocket.

The binding affinity of F103A for IBT was around 19 times weaker than the wild type. Although no NMR assignments were made for Phe 103 and its adjacent residues, this result points to IBT being in particularly close proximity to this residue and supports the evidence from the other mutants that IBT binds in this area of the pocket. This area is quite far removed from Y84 which is seen as being at the “top” or “open” end of the pocket and therefore raises the question of why Y84A makes such a large impact on binding, if it is possible that Y84 isn’t in contact with the ligand.

It has previously been mentioned in sections 1.4 and 5.2.8 that Y84 (equivalent to Y82 in OBP1 and porcine OBP) may play a role as a capping or gate-keeping residue which regulates entry into the ligand binding pocket (Meillour et al. 2009; White et al. 2009). The evidence for this is partly due to the location of the side chain which sits across the opening of the barrel; this is shown for OBP3 in figure 7.6. Additionally molecular dynamic simulations on porcine OBP had shown Tyr 82 and Phe 35 to be involved in ligand entry into the pocket, and subsequent studies of Y82A and F35A mutants had shown the fluorescent probe 1-AMA was not retained by the proteins unlike the wild type and so dissociation studies could not be conducted (Meillour et al. 2009). In OBP3 there is not a phenylalanine in the equivalent position to F35, however ITC has revealed

that the binding of the supposed cap mutant Y84A has drastically impaired binding.

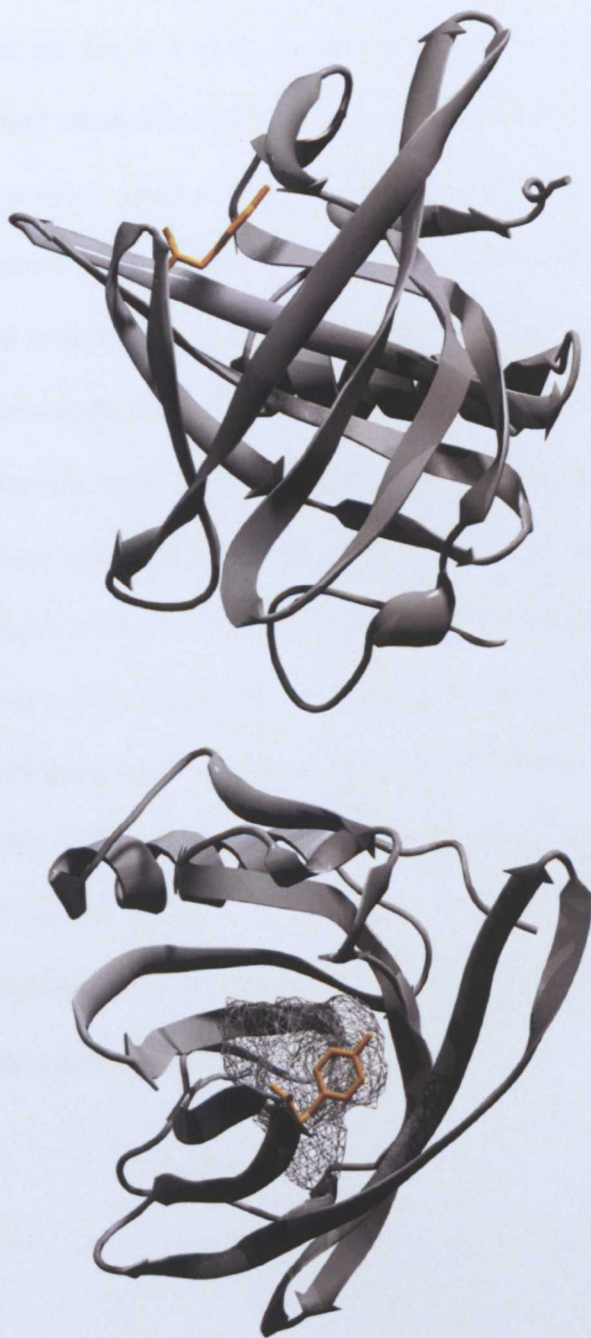


Figure 7.6 Tyr is highlighted in orange. The top panel shows OBP3 viewed with Tyr 84 at the top of the β -barrel. The way it forms an apparent “cap” over the binding (hatched) is best viewed in the bottom panel.

The thermodynamic parameters (shown in figure 7.7) of all the mutants revealed that all the reactions were enthalpically driven like the wild type protein; however the enthalpy and entropy changes were slightly smaller, with the exception of Y84A. Due to the nature of the Y84A binding curve (shown in figure 7.5) the parameters derived are not as reliable as the fit is more ambiguous and this is reflected in the larger standard errors calculated for this mutant. The interaction of IBT with Y84A is much more entropically unfavourable than with any of the other mutants, whilst the enthalpy change is more negative. If the pocket of Y84A is solvent exposed it may explain why the entropy is more unfavourable as the system is more disordered and, when the ligand attempts to bind, more order is being imposed upon the system rather than less (as would happen when ordered water molecules were released). The larger enthalpy term may point to the making and breaking of bonds within the protein as it attempts to compensate for the open pocket and continue to bind the ligand. It could also be due to altered interactions of the protein and ligand with the solvent. Despite the open pocket it is possible that the hydrophobic association between the binding cavity and the hydrophobic ligand was still strong enough for binding to occur. Additionally some electrostatic interactions involving the charged residues in the pocket, such as Asn 88 and Glu 118, may also have contributed to the interaction.

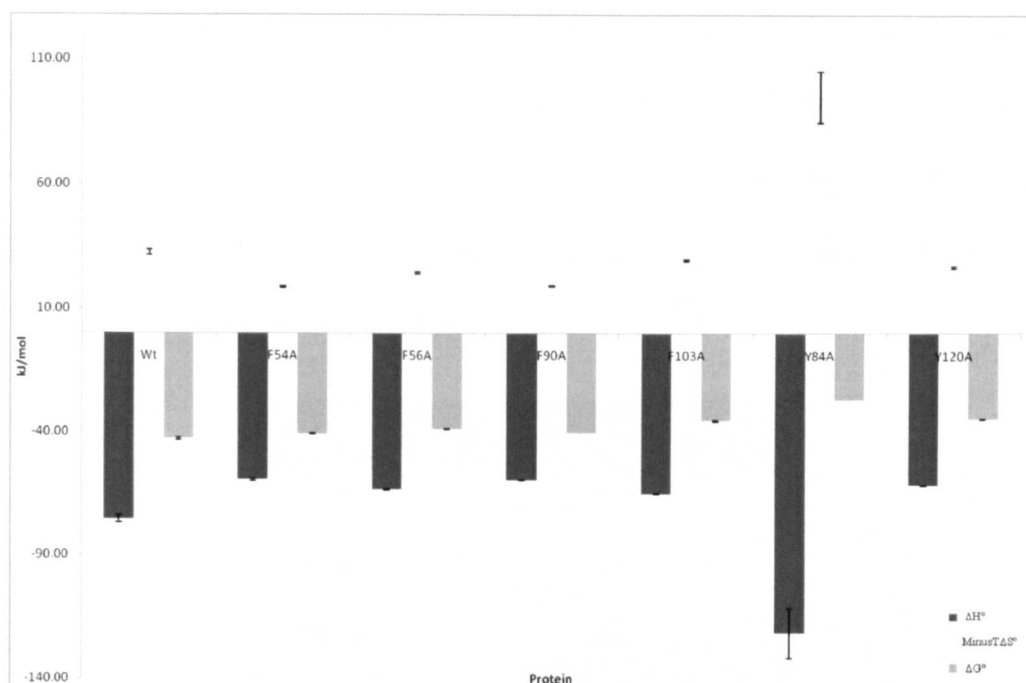


Figure 7.7 Thermodynamic parameters for each of the aromatic mutants binding to IBT. The binding of Y84A is markedly different. The error bars are the standard error of the mean of duplicate experiments.

To further establish the role of Tyr 84 in OBP3 an additional mutant was constructed, which changed the residue from a tyrosine to a phenylalanine. This mutant was made to determine if it was the phenyl ring of the residue that was important or if it was the hydroxyl group that was vital. The residue at this position is a tyrosine in OBP1 whilst it is a phenylalanine in OBP2. The binding profile of OBP2 is quite distinct from the other two rat OBPs (Löbel et al. 2002) which could be associated with the different residue at this position.

Y84F was produced in the same manner described for the other aromatic mutants. Mass spectrometry was used to verify the correct protein had been produced.

Additionally DSC was used to compare the stability of Y84F to Y84A and the wild type. As shown in table 7.2 a melting temperature of 344 K was determined, which was 1 K higher than Y84A and 2 K lower than the wild type. This indicated the protein was relatively stable (and importantly had very similar stability to Y84A).

Protein	Melting Temperature (K)
Wt OBP3	346
Y84A	343
Y84F	344

Table 7.2 Melting temperatures of wild type OBP3 and Tyr 84 mutants

An example of an ITC curve of Y84F binding to IBT, compared to the wild type protein is shown in figure 7.8. It is immediately obvious that the binding mode is similar to the wild type and the weak binding of Y84A is not seen. The average dissociation constant is 0.12 μM which, although 3 times weaker than the wild type protein, is still relative low. It is also weaker than F90A and F54A but shows similar affinity to F56A. This shows that although the hydroxyl group of the tyrosine may play a role in the binding of IBT it is not essential to maintain a strong affinity, and in fact it is the presence of the phenyl group that is essential.

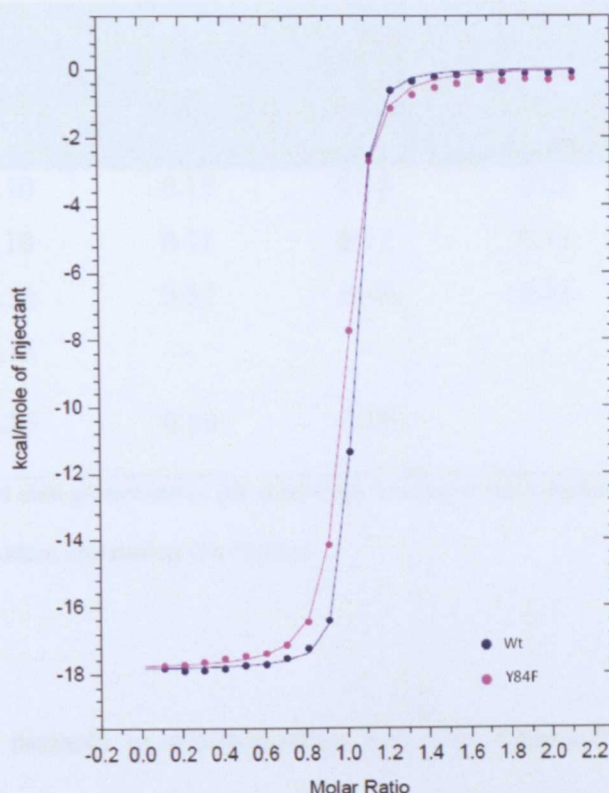


Figure 7.8 ITC binding curves of wild type OBP3 and Y84F binding to IBT. Overall the curves are very similar in nature.

7.2.3 ITC of the aromatic mutants with the γ -lactones

The chemical shift perturbations of OBP3 binding to the γ -lactones measured in section 6.2.2.2 were different from those measured for IBT, particularly for Phe 90 which was only significantly perturbed upon binding to the γ -lactones. The chemical shift perturbations for the aromatic residues mutated in this chapter are shown in table 7.3. Phe 103 is excluded as it was unassigned. Tyr 84 was not able to be assigned for any of the γ -lactones complexes except γ -heptalactone.

Residue	CSP upon binding γ - heptalactone	CSP upon binding γ - octalactone	CSP upon binding γ - nonalactone	CSP upon binding γ - decalactone	CSP upon binding γ - undecalactone
F54	0.10	0.15	0.18	0.23	-
F56	0.10	0.11	0.12	0.11	0.09
F90	0.30	0.32	0.40	0.41	0.36
Y84	0.15	-	-	-	-
Y120	0.25	0.26	0.29	-	-

Table 7.3 Chemical shift perturbations (in ppm) upon binding to the γ -lactones from section 6.2.4 for the aromatic residues mutated on this chapter.

Binding of the mutants to γ -heptalactone was only measurable for Y84F and Y120A, the ITC curves are shown in figure 7.9. Fitting the curve for Y84F was successful; the K_d derived (23 μ M) showed this mutant had a lower affinity for γ -heptalactone than Y84A had for IBT. It was not possible to satisfactorily fit the binding curve for Y120A so binding parameters could not be derived. Therefore γ -heptalactone was excluded from the mutant analysis as was γ -undecalactone which was lacking in CSP data for the majority of the mutated residues.

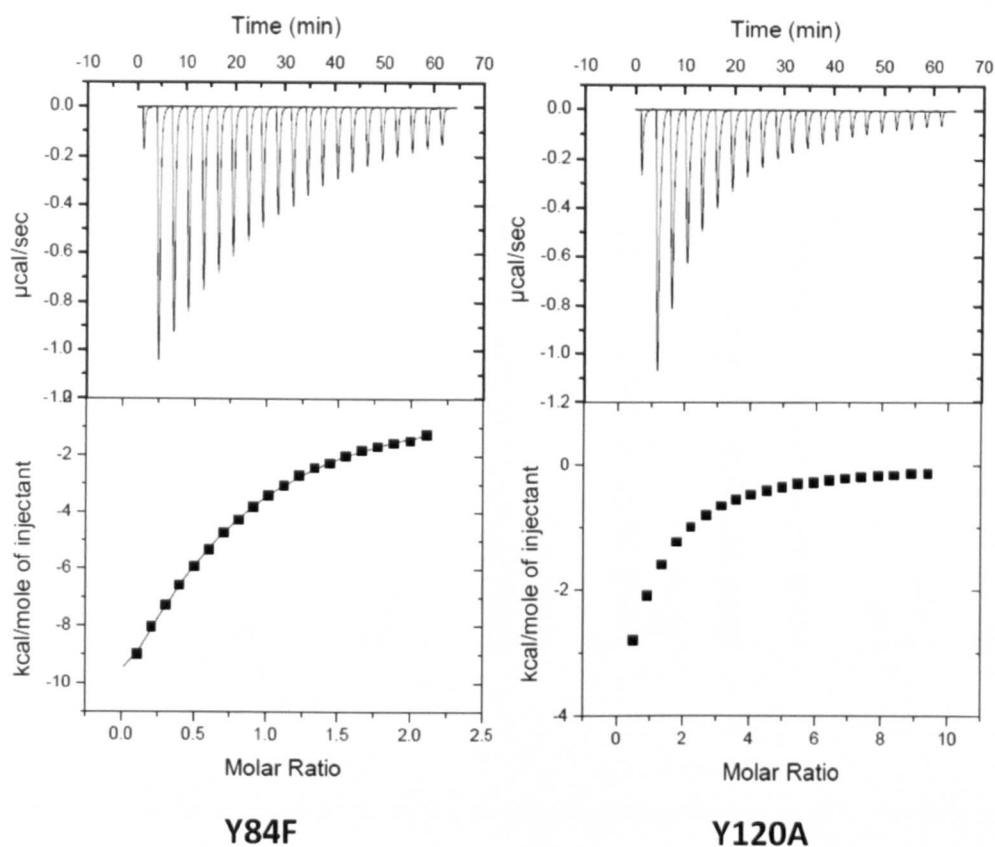


Figure 7.9 ITC binding curves for the only aromatic mutants that showed binding to γ -heptalactone ($K_d = 23 \mu\text{M}$, $\Delta H^\circ = 87.45 \text{ kJ/mol}$). A satisfactory fit for the binding curve of Y120A could not be achieved.

ITC experiments were carried out for each of the aromatic mutants with γ -octalactone, γ -nonalactone and γ -decalactone. The dissociation constants are plotted in figure 7.10. There are no values recorded for γ -octalactone and γ -nonalactone binding to Y84A, as a satisfactory fit could not be achieved. Also plotted on figure 7.10 are the values for each of the mutants binding to IBT for comparison. As with the wild type protein, binding of the γ -lactones is weaker for all but one mutant (with γ -decalactone) than with IBT.

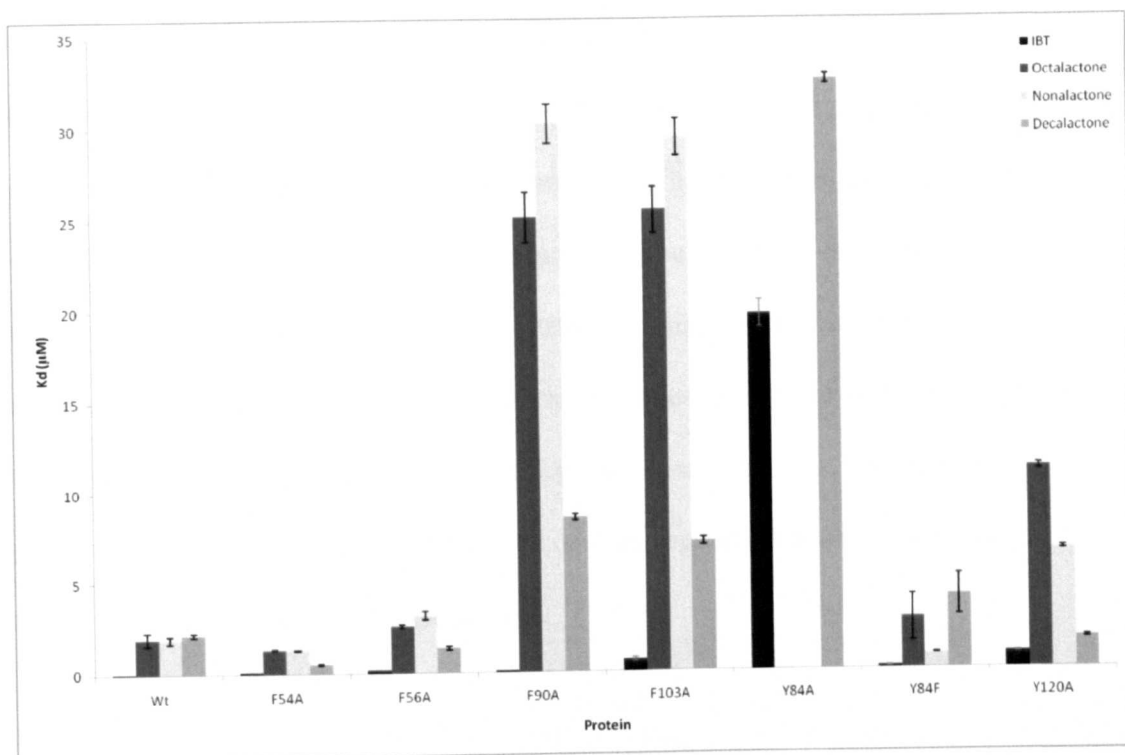


Figure 7.10 Dissociation constants for each of the aromatic mutants binding to IBT, γ -octalactone, γ -nonolactone and γ -decalactone. F90A, F103A and Y84A have markedly lower affinities than the other mutants. The error bars for the γ -lactones are standard errors of the mean of triplicate experiments, except for the wild type and IBT values which are the standard errors of the mean of duplicate experiments. Numerical values are given in table 7.4, along with the thermodynamic values.

The effect on affinity of the F54A mutant is interesting as, despite the chemical shift of Phe 54 being significantly perturbed to an increasing extent upon binding each of the γ -lactones; the mutation actually increases the affinity for each of the γ -lactones, with F54A having the greatest affinity for γ -decalactone. The affinity also remains high for F56A, although only the interaction with γ -decalactone has a higher affinity than the wild type. The shift of Phe 56 was perturbed to a similar, small, extent (0.10 ppm) upon binding to each of the γ -lactones. These results

suggest that in the wild type protein Phe54 and Phe 56 contribute to binding, but despite contributing to the hydrophobic interaction with the ligand, the side chain of Phe 54 may hinder the accommodation of the ligands. The increasing CSP value for Phe 54 seen with increasing ligand size would support the adjustment of the position of this residue to accommodate the longer hydrocarbon chain. The fact that Phe 56, which is very close to Phe 54, does not appear to have a large CSP contribution may suggest that the ligand doesn't occupy this part of the pocket and small effects on dissociation constant seen are a result of the pocket changing shape to allow the ligand to bind.

The chemical shift of Phe 90 was increasingly perturbed as each of the three ligands used for ITC were bound, this however doesn't relate to a similar trend in binding affinities. Although all the γ -lactones tested bound to F90A the affinity for γ -decalactone was three times higher than for γ -nonalactone. These results affirm the finding that Phe 90 is not involved in the binding of IBT. This is shown by a very small NH chemical shift perturbation of Phe 90 (table 7.1) and high affinity binding of F90A to IBT seen in ITC studies (figure 7.6) but is important in the binding of the γ -lactones. As Phe 90 is on the opposite side of the cavity to Phe 54 and Phe 56 it suggests the γ -lactones are accommodated near Phe 90 and Phe 90 is involved in hydrophobic interactions with the ligand. A possible explanation for the improved binding of F90A to γ -decalactone over γ -nonalactone and γ -octalactone is two-fold: firstly the lack of the aromatic side chain in F90A may allow more room in the pocket for the γ -decalactone hydrocarbon chain and secondly because of its longer hydrocarbon chain γ -decalactone is more hydrophobic and therefore more able to negate the reduced

hydrophobicity of the ligand pocket. In fact all the mutants, except Y84A and Y84F have higher binding affinity for γ -decalactone than the other γ -lactones, probably for similar reasons. The higher affinity F90A shows for this ligand supports the idea that the longer hydrocarbon side chain makes it more difficult to accommodate the larger ligand, it may be the case that in the wild type protein the ligand is shifted into a different position.

CSP values were not available for Phe 103; however the mutant binding affinities show a similar pattern and magnitude to Phe 90A, which can probably be explained in a similar way. Although the binding affinities were reproducible across the replicates of the experiments with γ -octalactone, the fits produced were not very similar and therefore this variation caused large error bars for the ΔH° parameter and these data should be treated with caution. Phe 103 is on the same side of the binding pocket as Phe 90, although it is further from Tyr 84 and the “top”. This again would suggest that the γ -lactones are accommodated by this side of pocket more than the side where Phe 54 and Phe 56 are.

Y120A showed a decreased affinity compared to the wild type for γ -octalactone (six times weaker) and γ -nonalactone (three times weaker). However, the K_d value for γ -decalactone was 1.69 μ M whilst the wild type value was weaker (2.12 μ M), (i.e. the removal of the side chain led to a higher binding affinity for the largest ligand than the wild type). This suggests the size of Tyr 120 causes problems accommodating the longer γ -lactones. This is supported by the fact that as the ligand size increased the CSPs for Tyr 120 became larger and the het-NOE values

for Leu116 and Tyr 120 (the residues in between were unassigned) (section 6.2.5) showed it was possible this area was becoming more rigid. The fact that the binding to γ -octalactone and γ -nonalactone were weaker than the wild type points to the fact that Tyr 120 makes some hydrophobic contribution to binding, though it doesn't necessarily point to Tyr 120 being involved in hydrogen bonding, either to the ligand or via solvent water.

Although changing residue 84 from a tyrosine to a phenylalanine had only a small impact on IBT binding, the effect on γ -lactone binding was greater. It is curious that γ -nonalactone was the least affected whilst the larger and smaller ligands were more affected (although it should be noted that γ -octalactone binding to Y84F was variable with a large standard error). This may suggest a more involved role for the tyrosine hydroxyl when binding to the γ -lactones. A greater involvement of this residue is also suggested by the fact that all the γ -lactone bound forms (except γ -heptalactone) were unassignable due to shifting to positions that were not able to be tracked. Additionally, the binding curves for γ -octalactone and γ -nonalactone binding to Y84A showed such weak interactions that the binding curves could not be fitted and hence the values are missing. The greater effect of the Y84A mutation seen on the binding of lactones than IBT is probably explained by the fact that IBT is a much higher affinity ligand, possibly through additional contacts being made, allowing it to be retained for longer in the pocket. The larger size of γ -decalactone may enable it to be retained to a certain degree due to it being more hydrophobic and having more room in the mutant pocket. Alternatively the pocket may be able to "close" round this ligand, strengthening the interaction somewhat.

The thermodynamic parameters for binding of the γ -lactones to each of the aromatic mutants (and the wild type for comparison) are shown in table 7.4. The phenylalanine mutants are grouped together, followed by the tyrosine mutants. For the weaker interactions (F103A to γ -octalactone and Y84A to γ -decalactone) the ΔH° and $T\Delta S^\circ$ values are large (and have large standard errors) due to the weakness of the interaction leading to unusual binding curve fits. Looking at the remainder of the values all the reactions remain enthalpically driven in a similar manner to the wild type protein. However, the entropy-enthalpy compensation observed for the wild type is not seen, which is most likely due to the complication caused by γ -decalactone binding more optimally than the smaller γ -lactones. There is no large variation across the mutants, with all showing larger ΔH° and $T\Delta S^\circ$ changes than the wild type. The lack of variation (with the exclusion of Y84A and Y84F) is likely due to the mutants being very similar in nature (despite being in different locations within the pocket). All mutants have the effect of making the pocket larger and decreasing the hydrophobic nature of the pocket to some degree. From the values it is apparent that Y120A is not very different from the phenylalanine mutants. This is in contrast to the effect of mutating Tyr 84 to an alanine, which has a large impact.

Protein	Ligand	K_d ($\times 10^{-3}$)	ΔH° (kJ mol ⁻¹)	ΔG° (kJ mol ⁻¹)	$T\Delta S$ (kJ mol ⁻¹)
Wt	γ -octalactone	1.92 \pm 0.34	-64.54 \pm 10.1	-32.66 \pm 0.45	-31.88 \pm 9.65
	γ -nonalactone	1.87 \pm 0.19	-53.51 \pm 0.08	-32.7 \pm 0.26	-20.82 \pm 0.34
	γ -decalactone	2.12 \pm 0.1	-44.77 \pm 0.9	-32.4 \pm 0.22	-12.4 \pm 1.26
F54A	γ -octalactone	1.32 \pm 0.02	-93.46 \pm 0.06	-33.55 \pm 0.05	-59.82 \pm 0.10
	γ -nonalactone	1.29 \pm 0.02	-41.05 \pm 0.24	-33.59 \pm 0.03	-7.46 \pm 0.25
	γ -decalactone	0.49 \pm 0.06	-64.52 \pm 0.82	-36.30 \pm 0.29	-27.40 \pm 1.11
F56A	γ -octalactone	2.6 \pm 0.08	-108.52 \pm 1.04	-31.87 \pm 0.09	-77.51 \pm 1.14
	γ -nonalactone	3.19 \pm 0.21	-54.77 \pm 0.88	-31.37 \pm 0.17	-23.4 \pm 1.05
	γ -decalactone	1.37 \pm 0.08	-86.82 \pm 0.88	33.60 \pm 0.15	-52.47 \pm 1.03
F90A	γ -octalactone	25.1 \pm 1.36	-113.85 \pm 10.0	-26.26 \pm 0.17	-82.09 \pm 10.2
	γ -nonalactone	30.2 \pm 1.07	-50.48 \pm 2.09	-25.72 \pm 2.09	-24.76 \pm 2.17
	γ -decalactone	8.58 \pm 0.16	-73.14 \pm 1.06	-28.92 \pm 0.05	-43.17 \pm 1.08
F103A	γ -octalactone	25.5 \pm 1.22	-326.77 \pm 133	-26.21 \pm 0.15	-300.56 \pm 134
	γ -nonalactone	29.43 \pm 1.01	-74.43 \pm 3.61	-25.78 \pm 3.61	-48.65 \pm 3.68
	γ -decalactone	7.19 \pm 0.19	-90.35 \pm 1.91	-29.28 \pm 0.07	-62.94 \pm 1.98

Protein	Ligand	K_d ($\times 10^{-3}$)	ΔH° (kJ mol ⁻¹)	ΔG° (kJ mol ⁻¹)	$T\Delta S$ (kJ mol ⁻¹)
Y84A	γ -octalactone				
	γ -nonalactone				
	γ -decalactone	32.55 \pm 0.25	-690.99 \pm 57.95	-25.60 \pm 0.02	-665.38 \pm 57.93
Y84F	γ -octalactone	2.79 \pm 1.29	-65.79 \pm 46.67	-31.99 \pm 1.24	-66.69 \pm 32.08
	γ -nonalactone	0.79 \pm 0.01	-55.79 \pm 1.17	-34.81 \pm 1.17	-20.98 \pm 1.15
	γ -decalactone	4.08 \pm 1.14	-54.98 \pm 4.77	-30.85 \pm 0.71	-24.13 \pm 5.48
Y120A	γ -octalactone	11.15 \pm 0.15	-70.28 \pm 43.05	-28.26 \pm 0.03	-77.16 \pm 1.56
	γ -nonalactone	6.63 \pm 0.08	-41.20 \pm 1.06	-29.55 \pm 0.03	-11.66 \pm 1.02
	γ -decalactone	1.69 \pm 0.07	-78.01 \pm 0.90	-32.93 \pm 0.10	-45.08 \pm 0.80

Table 7.4 The thermodynamic parameters for binding of each of the aromatic mutants to the γ -lactones. The errors stated are the standard error of the mean of 3 experiments with the exception of the wild type experiments which were conducted in duplicate.

The evidence points to the fact that Tyr 84 is the most important aromatic in the binding pocket, as whilst mutating of most other aromatics in the pocket leads only to subtle differences in binding affinities, removal of the aromatic side chain from Tyr 84 leads to drastic reduction, and in some cases, abolishing of binding. The precise cause of such a drastic change cannot be completely derived from the data, as the weaker binding leads to an inability to fit the ITC data and hence only a qualitative “lack of binding” can be ascribed. It has also been found that Phe 103 and Phe 90 are particularly crucial to the binding of the smaller γ -lactones, despite Phe 90 being apparently unimportant in the binding of IBT. The data on the aromatic mutants, as well as CSP and data on the residues reveals that the ligands tested do not all bind in the same position or area of the binding pocket. The data also show that IBT is able to compensate more readily than the γ -lactones for changes in the binding pocket. It may be the case that as well as generic hydrophobic interactions, IBT makes other specific interactions with residues in the binding pocket. The possibility of such specific interactions is explored further in the following section.

7.2.4 Design and production of specific binding pocket mutants

The affinity of IBT cannot be ascribed only to the hydrophobic interaction with the OBP3 binding pocket. As mentioned in section 6.2.5 IBT is not as hydrophobic as γ -decalactone, and its smaller size means it occupies a different part of the pocket, away from Phe 90 (hence this residue is shown to be particularly unperturbed in NMR studies) showing it to make fewer hydrophobic contacts. The thiazole ring contains a nitrogen atom which has the potential to form hydrogen bonds within the pocket. It was therefore interesting to explore some of the other residues in the OBP3

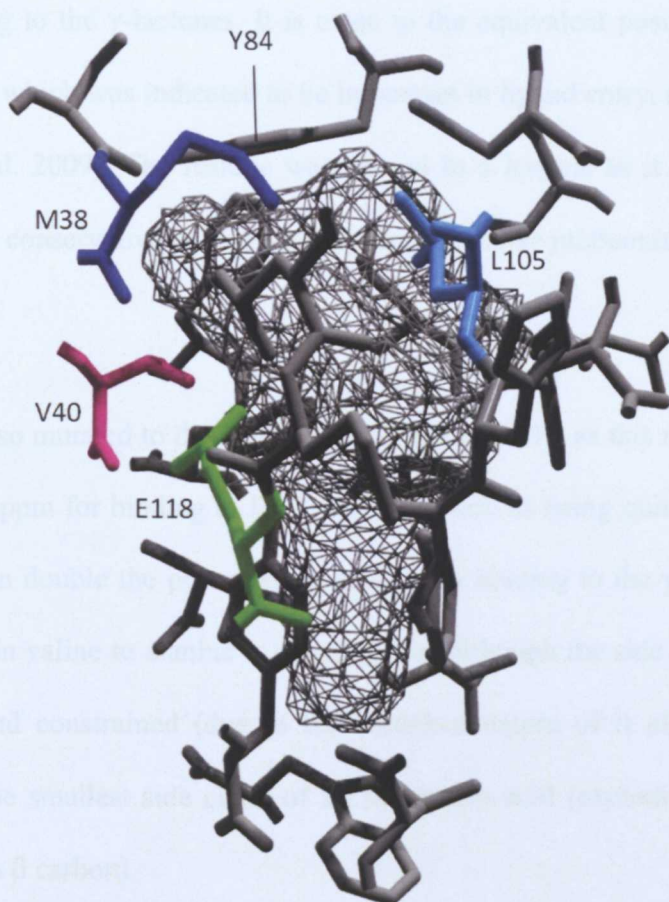


Figure 7.12 The specific residues targeted for mutation in this section. Tyr 84 denoted where the top of the pocket is.

The mutants constructed are highlighted by being in bold and underlined, with the mutant name being written alongside. In figure 7.12 the location of the residues targeted for mutation are highlighted on the structure. Tyr 84 is shown to highlight where the “top” of the pocket is.

Met 38 has a hydrophobic side chain, the residue was of interest however as it’s NH chemical shift was only significantly perturbed (0.07 ppm) upon binding IBT and

not by binding to the γ -lactones. It is close to the equivalent position of Phe 35 in porcine OBP, which was indicated to be important in ligand entry, along with Tyr 84 (Meillour et al. 2009). The residue was altered to a leucine as it is in OBP1. This substitution is conservative as leucine is hydrophobic like methionine.

Val 40 was also mutated to the equivalent residue in OBP1 as this residue had a CSP value of 0.18 ppm for binding to IBT, which, as well as being quite significant, was also more than double the perturbation seen when binding to the γ -lactones. Again, the substitution valine to alanine is conservative, although the side chain of valine is quite large and constrained (due to the branched nature of its alkyl chain) whilst alanine has the smallest side chain of all the amino acids (excluding glycine which doesn't have a β carbon).

The binding pockets of the odour binding proteins contain very few polar amino acids (Spinelli et al. 1998; Meillour et al. 2009; White et al. 2009). Asn 88 is conserved across a number of odour binding proteins (Spinelli et al. 1998). Its NH chemical shift is slightly perturbed (CSP = 0.08 ppm) upon binding IBT and all the γ -lactones, suggesting it may play a role in binding. Asn 88 is the only shared polar residue between OBP1 and OBP3. In addition, OBP3 possesses a glutamate in position 118 (which is a non-polar leucine in OBP1) and OBP1 has an asparagine at position 105 (which is a leucine in OBP3). Glutamate is a negatively charged amino acid and therefore it is possible it is involved in electrostatic interactions. The residue is also capable of interacting with positive amino acids and forming stabilising hydrogen bonds. In changing this residue to an alanine this negative charge will be

lost, as will the relatively large side chain of Glutamate. Glutamate 118 has been shown to be involved in hydrogen bonding that is important to the binding of 2-sec-butyl-4,5-dihydrothiazole (SBT) to nasal MUP, as discussed in section 5.2.6 (Perez-Miller et al., 2010) however this position is not conserved in other MUPs or OBPs (Spinelli et al. 1998; Perez-Miller et al. 2010). Glu 118 was not assigned by NMR but the nearby residues 113-116 in strand H showed indications of being important in accommodating ligands (section 5.2.8 and 6.2.4).

Asn 105 is conserved across a number of OBPs (Spinelli et al. 1998) despite this it is a leucine in OBP3. The L105N mutation introduces a polar, non-charged group into this position. Asn has the ability to hydrogen bond to the protein backbone, and so is often involved in α -helices; however it forms part of β -strand H in OBP3. Again this residue was unassigned by NMR, the closest assigned residues were marginally perturbed on binding ligands so the participation of Leu 105 in binding is difficult to ascertain. The side chain of asparagine is larger than leucine so this may have an influence on the binding of ligands as it makes the pocket smaller. Asparagine is often found in binding sites of proteins interacting with other polar groups (on other amino acids or the ligand) and therefore this may occur in L105N.

The binding pocket of the L105N mutant contains both an asparagine and a glutamate side chain, which may affect the pocket size. To help to avoid any steric problems caused by this, and to make the pocket more similar to the OBP1 pocket, a double mutant was constructed (L105N/E118L). The predicted effect of the mutations on pocket size was explored using Swiss PDB viewer (Guex & Peitsch

1997). The mutations were introduced and the molecular surface and cavity size calculated using the inbuilt functions in the Swiss PDB Viewer software package (Guex & Peitsch 1997). Interestingly all of the mutations appeared to make the volume of the pocket larger. This may be because when the mutated residue is modelled the most energetically favourable rotamer of the amino acid is selected and this may not necessarily be the case for the wild type protein in solution. The sizes of the modelled pockets are given in table 7.5.

Protein	Modelled pocket size (Å ³)
Wt OBP3	428
E118L	453
L105N	441
L105N / E118L	466

Table 7.5 The modelled pocket sizes of Glu 118 and Leu 105 mutants.

A comparison of the models for wild type OBP3 and the double mutant is shown in figure 7.13 and demonstrates that the overall pocket shape is quite similar, although the double mutant appears to be more expanded. His 104 has been excluded from the figure to make it easier to view the residues.

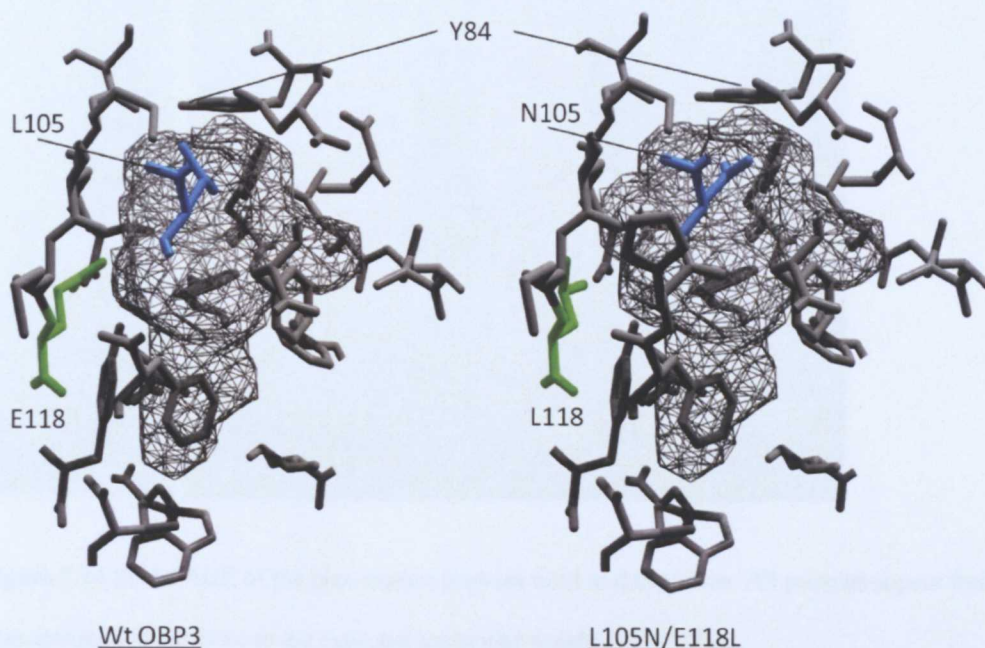


Figure 7.13 Comparison of the modelled binding pockets of wild type OBP3 and the L105N/E118L mutant. The double mutant is calculated to have a 38 Å³ larger pocket, though the overall shape is quite similar.

The same method used to produce the aromatic mutants (section 7.2.1) was also used to produce the these mutants. The primers used are listed in appendix I. The double mutant was constructed by using the E118L mutant plasmid as a template and performing the mutagenesis using the same primers used for L105N. All of the mutants were verified using SDS-PAGE to check the purity of the proteins and ESI-TOF mass spectrometry to verify that the correct mutation had been introduced. The SD-PAGE gels are shown in figure 7.14.

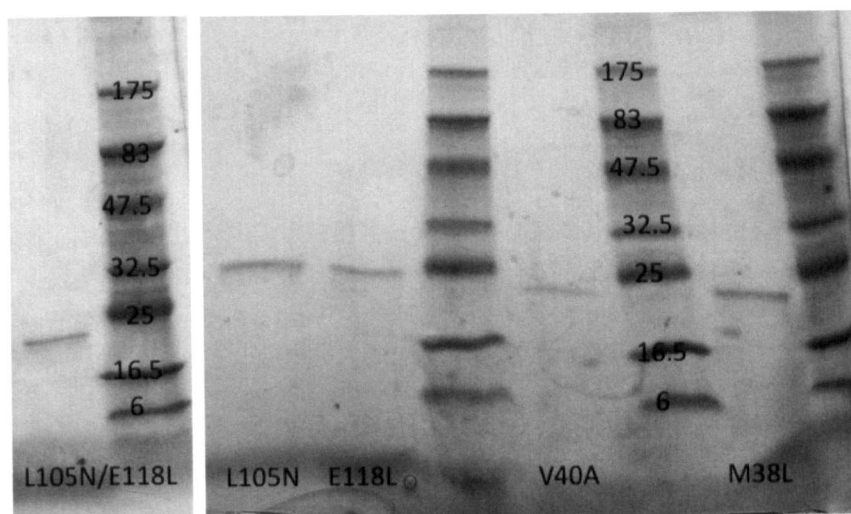


Figure 7.14 SDS-PAGE of the pure mutant proteins used in this section. All proteins appear free from contaminants and are close to the expected molecular weight (~19.9 kDa)

7.2.5 ITC of the specific mutants with the 2-isobutylthiazole

OBP1 has been shown to bind to 2-isobutylthiazole (IBT) with approximately two times weaker affinity than OBP3 (Löbel et al. 2002). To judge the role of each of the residues in the binding of IBT, binding of each of the mutants was analysed by ITC, conducted in the same manner as all previous ITC experiments described. The resulting affinities for each of the mutants (and the wild type for comparison) are shown in figure 7.15.

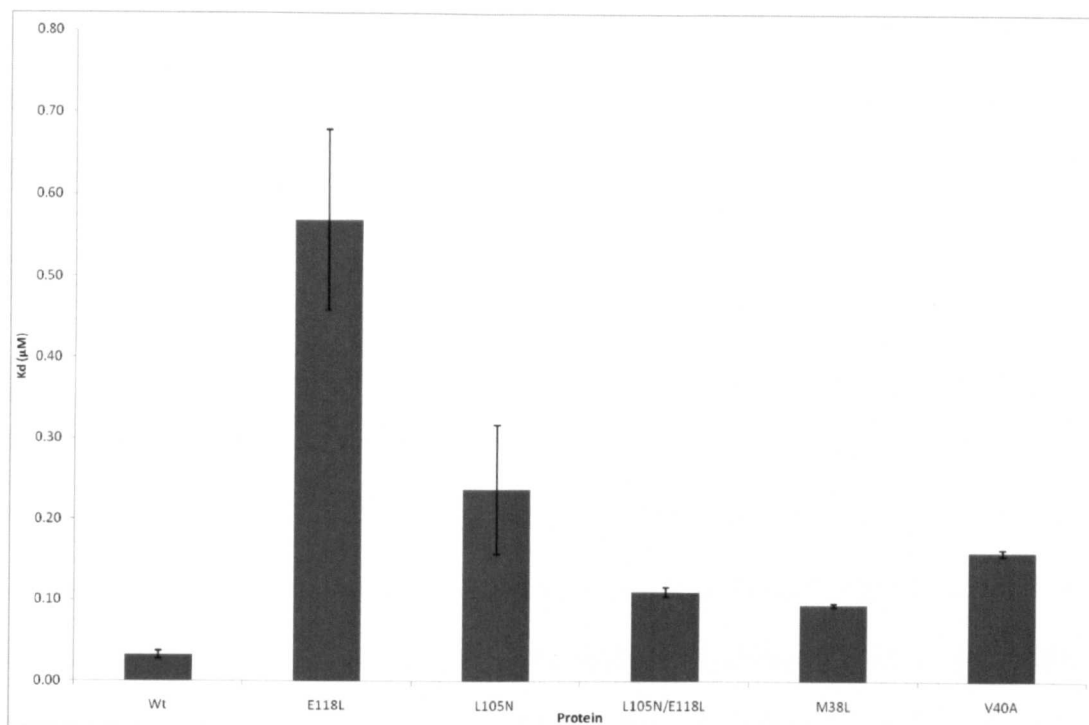


Figure 7.15 Binding affinities for each of the mutants binding to IBT. All mutants have a lower affinity for IBT than the wild type protein. The error bars are standard errors of the mean of triplicate experiments.

In a similar manner to the aromatic mutants (with the exclusion of Y84A) these mutants maintain high affinity interactions with IBT, with all K_d values being lower than 1 μ M. This means that for all these mutants IBT still has almost twice the binding affinity of wild type OBP3 binding to the γ -lactones. The smallest effect seen, as was expected from the NMR CSP values, was the M38L mutation. The substitution as mentioned previously, is quite conservative with both methionine and leucine being hydrophobic residues. Methionine does possess a sulphur atom in its side chain, however this is unable to form disulphide bonds as it is attached to a methyl group (rather than just a hydrogen atom in cysteine), so the role of methionine is usually to increase hydrophobic interactions. The fact that the mutation

gives some significant effect (binding is 3 times weaker than the wild type) shows that in OBP3 this residue is more optimal for IBT binding than the equivalent Leucine in OBP1.

The chemical shift of the amide proton of Val 40 was more perturbed on ligand binding than Met 38 and the weaker binding of IBT by V40A reflects this. V40A binds IBT five times more weakly than the wild type protein. The mutation introduces a smaller side chain. The difference in hydrophobicity between valine and alanine is relatively small, depending on the scale used (see section 7.1.1). From the affinity result, it appears that a valine at this position is more optimal than alanine for binding IBT. Val 40 is deeper in the pocket (further way from Tyr 84) than Met 38, which may indicate IBT interacts further into the pocket. However the binding affinity is the same as seen for F56A, which was concluded not to be highly important in IBT binding, so role of the Val 40 should not be over interpreted as the change in affinity is still relatively small.

The impact on affinity for IBT of the E118L mutation ($K_d = 0.57 \mu\text{M}$) is the largest of any of the specific mutants (19 times weaker than the wild type). It also had a larger impact than most of the aromatic mutations with the exception of Y84A ($K_d = 19.76 \mu\text{M}$) and Y120A ($K_d = 0.84 \mu\text{M}$). Taken together with the x-ray crystallography evidence on nasal MUP (Perez-Miller et al., 2010) this suggests that Glu 118 is likely to form a salt bridge with the nitrogen of IBT or possibly a hydrogen bond with Tyr 120. When this interaction is no longer present, however, the protein is still able to bind IBT, due to the presence of a large number of

hydrophobic interactions and the potential of Tyr 120 to hydrogen bond to the ligand. It is difficult to discern the contribution the change in size has to the binding affinity. It is possible that the larger pocket leads to IBT not being as close to the residue, but using ITC this is not distinguishable from the loss of possible charge interactions.

This is also true for L105N which shows a decreased binding affinity for IBT compared to the wild type (eight times weaker), however the cause of the effect is difficult to separate out. The introduction of the polar asparagine side chain may add to the binding of IBT; however the change in pocket size may cause the binding to become weaker. Additionally the reduction in hydrophobicity of the residue may also have lead to weaker binding. The double mutant, which removes the possible steric hindrance of having both Glu 118 and Asn 105 in the pocket, and is the most similar to the situation in OBP1, has a greater binding affinity than either of the single mutants, however the binding is still three and half times weaker than the wild type. This is also weaker affinity than OBP1 demonstrates for IBT (Löbel, 2002). Of course the remainder of the binding pocket is not identical to OBP1 and therefore exactly the same binding cannot be expected. It is interesting to note however that despite a possible missing electrostatic interaction provided by E118 the double mutant is able to bind well to IBT.

The thermodynamic parameters shown in figure 7.16 reveal that all of the mutants show a (similar) enthalpically driven mode of binding to the wild type protein. The entropy is displayed as $-T\Delta S^\circ$ to enable it to be distinguished more readily. M38L and V40A, being the most conservative mutations, show enthalpic and entropic

contributions to binding most similar to the wild type protein. The introduction of an asparagine at position 105 leads to a smaller entropic penalty than either the wild type protein or any of the other mutants, which may be a reflection of the introduction of an additional polar group into the pocket, this is also seen in the double mutant to a lesser extent as it may be negated by the loss of the glutamate at position 118.

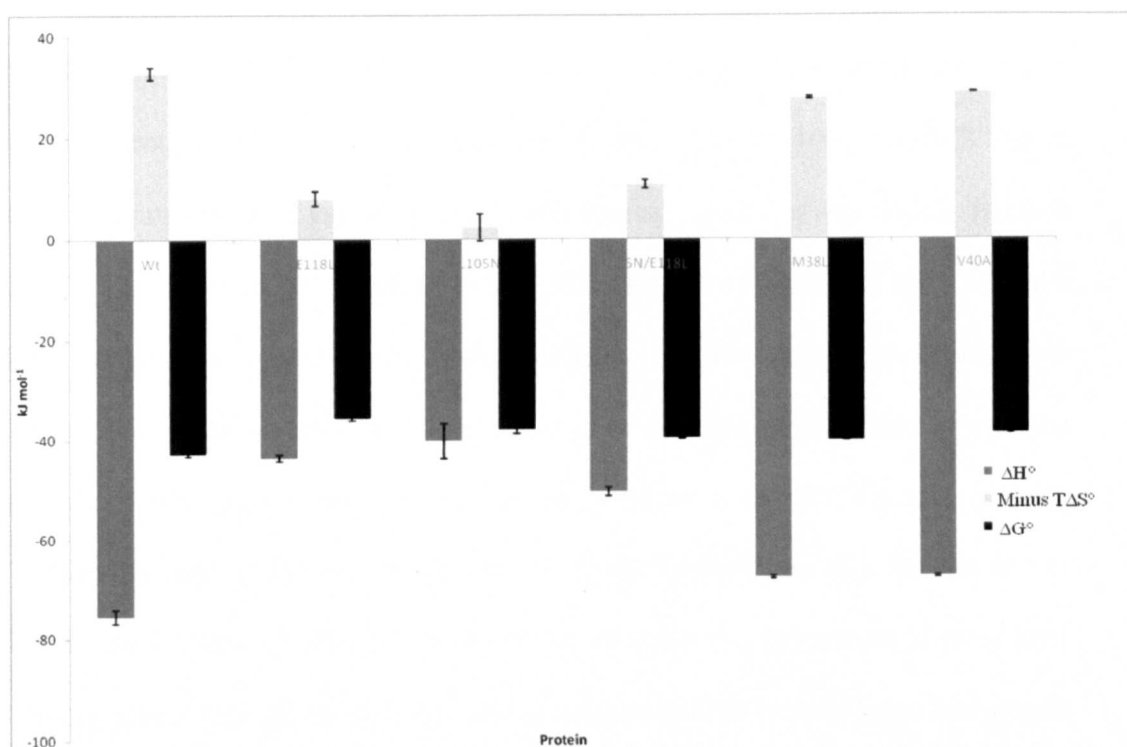


Figure 7.16 Thermodynamic parameters for each of the specific mutants binding to IBT. The error bars are the standard error of the mean of triplicate experiments.

Due to the particularly high affinity binding of OBP3 to IBT it is unsurprising that none of the mutants improved the affinity. The data on the specific mutants, along with the aromatic mutants, highlight how apparently optimised OBP3 is for binding

to IBT, with every residue tested making some contribution to the high affinity. OBP1 however does bind more strongly than OBP3 to some ligands, for example the acetate esters (6.2.6), it was therefore of interest to investigate whether some of the mutants were able to enhance the binding of OBP3 to sec butyl acetate and tert butyl acetate.

7.2.6 ITC of the specific mutants with the acetate esters

Sec butyl acetate (SBA) and tert-butyl acetate (TBA) were selected rather than isobutyl acetate (IBA) as they showed particularly weak binding to OBP3. Again ITC experiments were carried out in same manner described previously (section 5.2.2). The E118L, L105N and L015N/E118L mutants were used as they were the least conservative mutations and were capable of more than just hydrophobic interactions. Unfortunately no interaction was seen between TBA and any of the mutants, which suggests that the ligand is not retained in the OBP3 binding pocket. The arrangement of this isomer appears not to allow interaction with residue 105 or 118 in the mutants of OBP3. It may be the case that the remainder of the OBP1 binding pocket residues or the shape and smaller size of the pocket helps this protein to retain TBA.

SBA was able to bind to wild type OBP3, but not as tightly as OBP1. The dissociation constants for this ligand to the OBP3 mutants (and the wild type for comparison) are shown in figure 7.17. The K_d values for the mutants binding to IBT are also included.

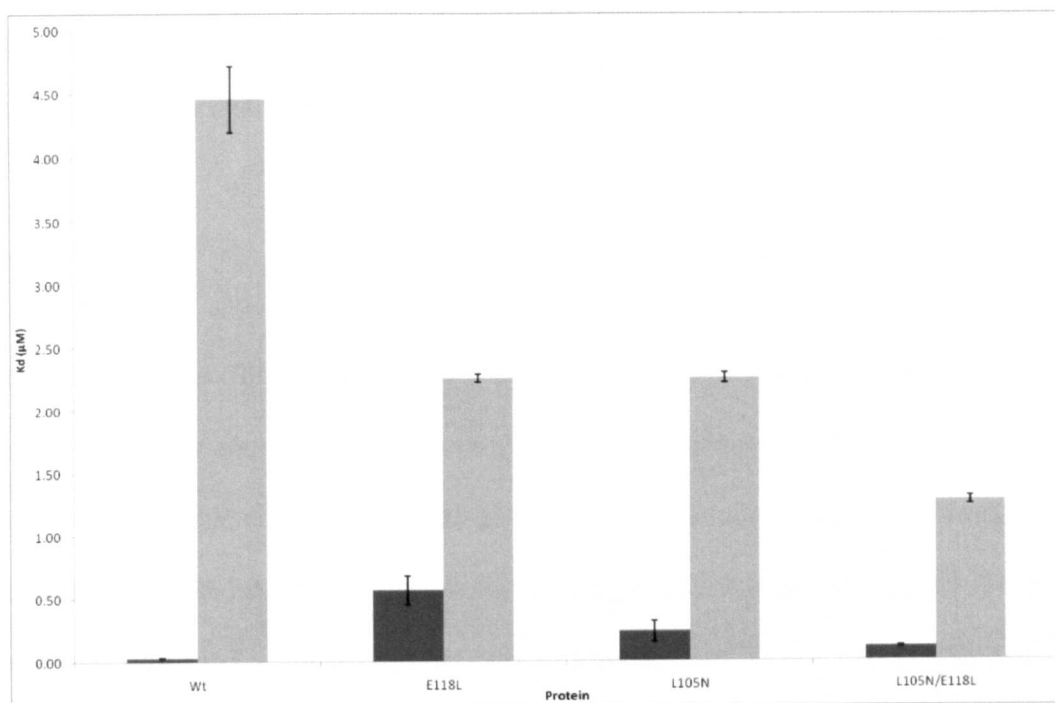


Figure 7.17 Dissociation constants for each of the specific mutants binding to SBA (light grey). The dissociation constants of the same mutants binding to IBT are shown in dark grey for reference. All the mutants have a higher affinity for SBA than the wild type protein. The error bars are standard errors of the mean of triplicate experiments.

The data clearly show that all of the OBP3 mutants improve the binding affinity to SBA. In a similar manner to the binding of IBT the double mutant shows higher affinity than the single mutants. In contrast to the binding of IBT, the single mutants each show equal binding affinity for SBA. This indicates that when wild type OBP3 binds to SBA Glu 118 hinders the binding somewhat. This is possibly due to the negative charge of Glu 118 being unfavourable or the large side chain causing a problem. Despite this residue still being present, L105N is able to improve the binding to SBA and indicates that the asparagine has a positive effect upon the interaction. In the double mutant the favourable asparagine is present and the

unfavourable glutamate is removed which leads to markedly increased binding affinity (3.5 times stronger than the wild type).

The enthalpy and entropy values (figure 7.18) show the positive entropic contribution seen for wild type OBP3 is lost for the single mutants; however there is a slightly larger enthalpic contribution which may indicate an additional structurally determined component to the binding, perhaps due to additional interactions. The double mutant shows both a larger favourable enthalpic contribution than E118L and the wild type, and a lower entropic cost to binding than E118L and L105N which together appear to improve the binding of the double mutant the most.

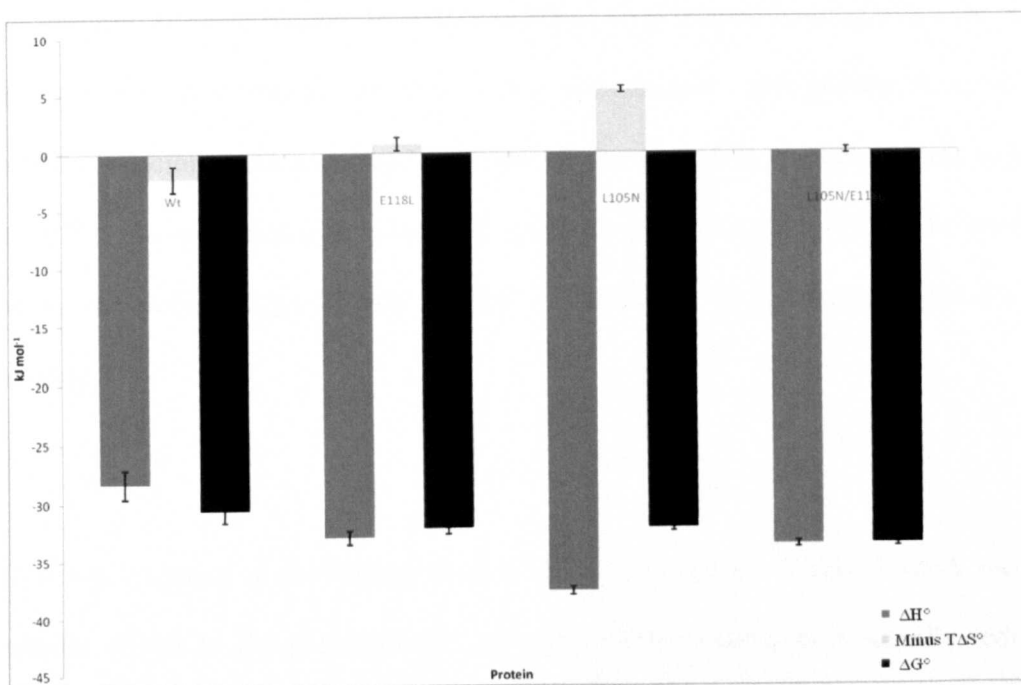


Figure 7.18 Thermodynamic parameters for each of the specific mutants binding to SBA. The error bars are the standard error of the mean of triplicate experiments.

The binding of the double mutant of OBP3 approaches the affinity of OBP1 for SBA, however it is clear that other aspects of OBP1 contribute to binding as the interaction is still around two times stronger (Borysik et al., 2010). The data are useful for showing that the affinities of OBP3 can be logically tuned to improve the affinity of certain ligands. They also point to a source of difference in the binding profiles of OBP1 and OBP3 and do indeed suggest that the nature of the binding pocket plays an important role in determining which ligands bind well.

7.3 Conclusions

The result of mutating a number of the residues in the OBP3 has been investigated. The conserved phenylalanine residues were substituted with alanine residues resulting in reduced binding to IBT. The effect on γ -lactone binding was even more pronounced. Interestingly the affinity for γ -decalactone was greater than for the smaller γ -lactones. As suggested in chapter 6 the size of the pocket appears to have an effect on the binding to γ -lactones, particularly the larger ligands. The mutants make the pocket bigger which enables γ -decalactone to be accommodated more readily.

The two tyrosines in the binding pocket were also mutated. Whilst Y120A showed similar effects to the phenylalanine mutants, Y84A resulted in drastically reduced binding affinities for IBT and the γ -lactones. A more conservative mutant, Y84F, did not show such a large loss of binding affinity and was comparable to the other mutants. DSC showed that neither of the Tyr 84 mutants had a large effect on the

protein stability. These mutants demonstrate the importance of a phenyl side chain at this position.

Selective mutants demonstrated that the residues in the OBP3 binding pocket are particularly optimal for 2-isobutylthiazole binding. The conservative hydrophobic mutations (M38L and V40A) showed the smallest reduction in binding affinity. Logically changing E118 and L105 to their equivalent residues in OBP1, as single mutants (E118L and L105N) and as a double mutant (E118L / L105N) had a detrimental effect on binding to IBT. However a ligand with high affinity binding to OBP1, sec-butyl acetate, showed improved binding to all three of these mutants. This effect was seen despite the fact that the mutants also increased the size of the binding pocket. SBA is smaller than IBT and therefore a larger pocket may be expected to have a negative effect on binding.

8. DISCUSSION AND FUTURE WORK

Within this thesis the binding profile of Odour Binding Protein 3 has been investigated to discover more about the nature of the binding interactions of vertebrate OBPs. OBPs are evolutionary conserved across vertebrates and invertebrates indicating their importance. OBPs in insects have been more fully studied than vertebrate OBPs and therefore their role is better understood. Vertebrate OBPs have been isolated from olfactory mucus in high concentrations and suggested to be transport proteins. Rat OBPs are particularly important as they have multiple subtypes with divergent sequences and previous studies have suggested that they interact with separate sets of chemically divergent odours.

In this thesis ITC was used to look more closely at OBP3 which appeared to be the most promiscuous rat OBP. In concurrence with previous studies heterocyclic compounds were found to be the highest affinity ligands of OBP3. Dissociation constants were found to be in the nanomolar range for these compounds, rather than the micromolar range found for the other odours investigated in this study. Whilst heterocyclic odour binding was entirely enthalpically driven, the binding of odours containing hydroxyl groups had a significant entropic contribution.

Very few vertebrate OBPs have been structurally characterised. It was therefore useful to attempt to structurally characterise OBP3. A number of MUPs, with relatively high sequence identity to OBP3 have had their high resolution structure solved, so it was anticipated the x-ray crystal structure of OBP3 could be generated.

Crystal production was the major obstacle that prevented this; however it is possible that further experimentation with the crystallisation conditions could result in diffractable OBP3 crystals. Saturating OBP3 with an odour prior to crystallisation attempts may also be a route to success.

NMR could also be used to solve the structure of OBP3. This would be more easily achieved with a more complete assignment of OBP3 (particularly β -strand G). Assignment of this area was hampered by missing peaks in the 3D spectra collected; meaning connectivities could not be established. Recording spectra at a different pH might make these peaks observable and it may also be useful to use a spectrometer at a higher frequency than 600 MHz. Experiments which directly recorded carbon resonances proved particularly useful for assigning OBP3 and “protonless” experiments which correlate carbon and nitrogen frequencies may help to establish the missing connectivities. In order to enable the residues involved in ligand binding to be visualised, and the location of the binding pocket to be estimated a homology model of OBP3 was constructed based on MUP I bound to sec butyl thiazoline. This model was also used to assess the effect of binding pocket mutants on the overall pocket size

NMR titrations revealed that binding to 2-isobutylthiazole gave NH resonances in slow exchange meaning binding constants could not be directly measured. The chemical shift perturbations revealed that ligand binding to OBP3 by both IBT and the γ -lactones did not cause large movements of the structure of OBP3, with all values recorded being below 0.5 ppm. Instead moderate adjustments were seen,

particularly of residues within the predicted binding pocket, to allow the accommodation of the ligands. $^1\text{H} / ^{15}\text{N}$ heteronuclear NOE values showed the majority of the residues in OBP3 to be quite rigid. The major exceptions were Gln 159 and Ala 160 at the C-terminus which were flexible and also the N-terminal residues preceding the first β -strand, which were more flexible than the remainder of the protein.

Across the γ -lactone series used there were general similarities in the residues that showed NH chemical shift perturbations; however there were differences between the binding of IBT and the γ -lactones. Most notably the NH shift of Phe 90 in the binding pocket was greatly perturbed by γ -lactone binding but was unchanged when IBT was bound to OBP3. It seems the odours occupy the binding pocket in subtly different ways. The reasons behind the particularly higher affinity binding of IBT are not immediately obvious from these experiments, although the involvement of Tyr 120 in hydrogen bonds appears possible from the CSP data. It would be interesting to undertake odour competition experiments to see if, for example, any of the γ -lactones were able to bind to OBP3 pre-saturated with IBT, particularly as biomimetic studies have shown the ability of IBT to displace other odours (Yabuki et al., 2011). This would have implications on which interactions are physiologically important. It may be the case that the γ -lactones cannot displace IBT and in the physiological setting the interaction doesn't occur.

By using a series of γ -lactones of increasing size possible entropy-enthalpy compensation was witnessed. The γ -lactones all had similar free energies of binding,

however the ΔH° and ΔS° values revealed this was achieved by apparent entropic compensation as the interactions of the larger ligands became less enthalpically driven. It was particularly interesting to see that adjustments within the OBP3 binding pocket (and some residues away from the binding pocket) were made to accommodate the large γ -lactones and maintain the binding affinity. $^1\text{H} / ^{15}\text{N}$ heteronuclear values of the γ -lactones bound to OBP3 showed this wasn't being achieved by an increased flexibility of large parts of the protein. Instead the protein remained relatively rigid whilst, for example, Leu 116 and nearby residues in β -strand H were shifted to allow the larger ligands to bind.

It is difficult to establish the orientation of the γ -lactones in the binding pocket, therefore solving the x-ray crystal structures of OBP3 bound to each of the lactones would be informative. Alternatively modelling the likely positions of these odours in the OBP3 binding pocket would be useful. In addition simulations of the entry and exit of, particularly the larger γ -lactones, would provide a lot of information,

The fact that γ -decalactone bound more tightly to the aromatic mutants of the OBP3 binding pocket than γ -octalactone or γ -nonalactone strongly suggests that the larger pocket produced by the mutants more comfortably accommodated this ligand. The loss of hydrophobic interactions and possible stacking interactions however did not, therefore, make the binding stronger than in the wild type protein.

The Y84A aromatic mutation dramatically reduced the binding affinity for IBT and abolished binding of the lower affinity γ -lactones. In previous studies on porcine OBP, no binding of the Y82A (which is equivalent to Y84A) to any ligand could be demonstrated as the fluorescent probe used for dissociation studies was unable to bind to the Y82A mutant. It is therefore useful to note that an interaction is possible with IBT, although the exact nature of this binding is undetermined. Crystallisation of this mutant may be interesting; however NMR would be particularly useful as dynamic measurements can be. The chemical shift perturbation data recorded in this study did not reveal a large change in shift of this residue, which suggests the unbound and bound positions of this residue are similar therefore more information could be gained from experiments to track the assumed dynamics of this residue when it allows an odour to bind (as shown by the modelling on porcine OBP).

Experiments on the acetate ester series revealed the role of ligand orientation to be important, the three isomers bound with particularly varied affinities, which was in contrast to previous studies on rat OBP1 which had similar affinities for these acetate esters. A particularly striking result was the ability of the E118L /L105N mutant of OBP3 to improve the binding of sec-butyl acetate and provides the first evidence that it may be possible to rationally “tune” the binding affinities of the OBPs to particular odours. This would be useful as this may enable these protein to be used to detect environmental pollutants, many of which are very similar to the odours OBP3 has been shown to bind. To explore this further it would be useful to look at more odours binding to this double mutant, and also to look at the interaction of E118L / L105N OBP3 with sec-butyl acetate by NMR. Constructing the “reverse” mutation

of OBP1 and monitoring acetate ester binding would also affirm the importance of these residues.

Despite the fact that the mutation of specific residues in the binding pocket of OBP3 are able to change the binding affinity of the protein the overall conclusion from this work is that OBP3 is indeed a promiscuous odour binder. It is able to accomodate a range of ligands with varying functional groups and sizes. Data from this study point to the fact that IBT is an ideal ligand due to a combination of hydrophobicity and the possible ability to form polar interactions. All mutants made to the OBP3 binding pocket had a detrimental effect on the affinity for IBT, making the pocket appear “optimised” for binding this ligand. However, whether this is the case in olfactory system is yet to be fully established.

REFERENCES

- Abbate, F., Franzoni, L., Löhr, F., Lücke, C., Ferrari, E., Robert, T. & Spisni, A. (1999) Complete ^1H , ^{15}N and ^{13}C assignment of a recombinant mouse major urinary protein. *Journal of Biomolecular NMR*, (15), pp.187-188.
- Al-Shawi, R., Ghazal, P., Clark, A J. & Bishop, J. (1989) Intraspecific evolution of a gene family coding for urinary proteins. *Journal of Molecular Evolution*, 29 (4), pp.302-313.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology*, 215 (3), pp.403-410.
- Amoia, A.. & Monfort, W.. (2007) Apo-nitrophorin 4 at atomic resolution. *Protein Science*, 16 (9), pp.2076-2081.
- Andersen, J.F., Weichsel, A., Balfour, C.A., Champagne, D.E. & Montfort, W.R. (1998) The crystal structure of nitrophorin 4 at 1.5 Å resolution: transport of nitric oxide by a lipocalin-based heme protein. *Structure*, 6 (10), pp.1315-1327.
- Arnold, K., Bordoli, L., Kopp, J. & Schwede, T. (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics (Oxford, England)*, 22 (2), pp.195-201.
- Ashkenazi, A., Presta, L.G., Marsters, S., Camerato, T.R., Rosenthal, K., Fendly, B.M. & Capon, D.J. (1990) Mapping the CD4 binding site for human immunodeficiency virus by alanine-scanning mutagenesis. *Proceedings of the National Academy of Sciences of the United States of America*, 87 (18), pp.7150-4.
- Atreya, H. & Chary, K. (2000) Amino acid selective “unlabelling” for residue-specific NMR assignments in proteins. *Current Science*, 79 (4), pp.13-16.
- Barratt, E., Bingham, R., Warner, D., Laughton, C., Phillips, S.E.. & Homans, S.W. (2005) Van der Waals interactions dominate ligand-protein association in a protein binding site occluded from solvent water. *Journal of the American Chemical Society*, 127 (33), pp.11827-34.
- Barratt, E., Bronowska, A., Vondrásek, J., Cerný, J., Bingham, R., Phillips, S. & Homans, S.W. (2006) Thermodynamic penalty arising from burial of a ligand polar group within a hydrophobic pocket of a protein receptor. *Journal of Molecular Biology*, 362 (5), pp.994-1003.
- Benkert, P., Biasini, M. & Schwede, T. (2011) Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics (Oxford, England)*, 27 (3), pp.343-50.
- Benton, R., Sachse, S., Michnick, S.W. & Voss hall, L.B. (2006) Atypical membrane topology and heteromeric function of *Drosophila* odorant receptors in vivo. *PLoS Biology*, 4 (2), pp.240-257.

- Bermel, W., Bertini, I., Duma, L., Felli, I., Emsley, L., Pierattelli, R. & Vasos, P.R. (2005) Complete assignment of heteronuclear protein resonances by protonless NMR spectroscopy. *Angewandte Chemie (International ed. in English)*, 44 (20), pp.3089-92.
- Bermel, W., Bertini, I., Felli, I.C. & Pierattelli, R. (2009) Speeding up ¹³C direct detection biomolecular NMR spectroscopy. *Journal of the American Chemical Society*, 131, pp.15339-15345.
- Bermel, W., Bertini, I., Felli, I.C., Kümmerle, R. & Pierattelli, R. (2006) Novel ¹³C direct detection experiments, including extension to the third dimension, to perform the complete assignment of proteins. *Journal of Magnetic Resonance*, 178 (1), pp.56-64.
- Betts, M.. & Russell, R.B. (2003) Amino acid properties and consequences of substitutions. In: M. . Barnes & I. C. Gray eds. *Bioinformatics for Geneticists*. Wiley, pp.289-318.
- Bianchet, M.A., Bains, G., Pelosi, P., Pevsner, J., Synder, S.H., Monaco, H.L. & Amzel, L.M. (1996) The three dimensional structure of bovine odorant binding protein and its mechanism of odor recognition. *Nature Structural Biology*, 3 (11), pp.934-939.
- Bignetti, E., Cavaggioni, A., Pelosi, P., Persaud, K.C., Sorbi, R.T. & Tirindelli, R. (1985) Purification and characterisation of an odorant-binding protein from cow nasal tissue. *European Journal of Biochemistry*, 149, pp.227-231.
- Bingham, R.J., Findlay, J.B.C., Hsieh, S.Y., Kalverda, A.P., Kjellberg, A., Perazzolo, C., Phillips, S.E.V., Seshadri, K., Trinh, C.H., Turnbull, W.B., Bodenhausen, G. & Homans, S.W. (2004) Thermodynamics of binding of 2-methoxy-3-isopropylpyrazine and 2-methoxy-3-isobutylpyrazine to the major urinary protein. *Journal of the American Chemical Society*, 126 (6), pp.1675-81.
- Bishop, J.O., Clark, A.J., Clissold, P.M., Hainey, S. & Francke, U. (1982) Two main groups of mouse major urinary protein genes, both largely located on chromosome 4. *EMBO Journal*, 1 (5), pp.615-620.
- Bocskai, Z., Groom, C.R., Flower, D.R., Wright, C.E., Phillips, S.E.V., Cavaggioni, A., Findlay, J.B.C. & North, A.C.T. (1992) Pheromone binding to two rodent urinary proteins revealed by X-ray crystallography. *Nature*, 360, pp.186-188.
- Borysik, A.J., Briand, L., Taylor, A.J. & Scott, D.J. (2010) Rapid odorant release in mammalian odour binding proteins facilitates their temporal coupling to odorant signals. *Journal of Molecular Biology*, 404 (3), pp.372-80.
- Bozza, T.C. & Mombaerts, P. (2001) Olfactory coding: revealing intrinsic representations of odors. *Current Biology*, 11 (17), pp.R687-90.

- Brechbühl, J., Klaey, M. & Broillet, M.-C. (2008) Grueneberg Ganglion Cells Mediate Alarm Pheromone Detection in Mice . *Science* , 321 (5892), pp.1092-1095.
- Breer, H. (2003) Olfactory receptors: molecular basis for recognition and discrimination of odors. *Analytical and Bioanalytical Chemistry*, 377 (3), pp.427-433.
- Breer, H. (2008) The sense of smell: reception of flavors. *Annals of the New York Academy of Sciences*, 1126, pp.1-6.
- Briand, L., Eloit, C., Nespoulous, C., Bezirard, V., Huet, J.C., Henry, C., Blon, F., Trotier, D. & Pernollet, J.C. (2002) Evidence of an odorant-binding protein in the human olfactory mucus: location, structural characterization, and odorant-binding properties. *Biochemistry*, 41, pp.7241-7252.
- Brown, R.C. (2007) γ -Lactones in wine : Synthesis , quantification and sensory studies. Flinders University of South Australia.
- Buck, L. & Axel, R. (1991) A Novel Multigene Family May Encode Odorant Receptors: A Molecular Basis for Odor Recognition. *Cell*, 65, pp.175-187.
- Burova, T.V., Choiset, Y., Jankowski, C.K. & Haertle, T. (1999) Conformational stability and binding properties of porcine odorant-binding protein. *Biochemistry*, 38, pp.15043-15051.
- Carey, A.F. & Carlson, J.R. (2011) Insect olfaction from model systems to disease control. *Proceedings of the National Academy of Sciences of the United States of America*, 108 (32), pp.12987-12995.
- Cavanagh, J., Fairbrother, W., Palmer III, A.G. & Sklepton, J. (2006) *Protein NMR Spectroscopy - Principles and Practice*. Academic Press.
- Chasman, D. (2003) Prospects for high-throughput structure determination of proteins by NMR spectroscopy. In: *Protein Structure: Determination, Analysis and Applications for Drug Discovery*. p.103.
- Chaudhuri, B.N., Kleywegt, G.J., Björkman, J., Lehman-McKeeman, L.D., Oliver, J.D. & Jones, T.A. (1999) The structures of α 2u-globulin and its complex with a hyaline droplet inducer. *Acta Crystallographica Section D Biological Crystallography*, 55 (4), pp.753-762.
- Chenna, R. (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Research*, 31 (13), pp.3497-3500.
- Cianci, M., Rizkallah, P.J., Olczak, A., Raftery, J., Chayen, N.E., Zagalsky, P.F. & Helliwell, J.R. (2002) The molecular basis of the coloration mechanism in lobster shell: beta-crustacyanin at 3.2-Å resolution. *Proceedings of the National Academy of Sciences of the United States of America*, 99 (15), pp.9795-800.

- Clark, J.A., Clissold, P.M., Shawi, R.A., Beattie, P. & Bishop, J. (1984) configurations in the 3' -non-coding region. *EMBO Journal*, 3 (5), pp.1045-1052.
- Clore, G.M., Driscoll, P.C., Wingfield, P.T. & Gronenborn, A.M. (1990) Analysis of the backbone dynamics of interleukin-1 beta using two-dimensional inverse detected heteronuclear ¹⁵N-¹H NMR spectroscopy. *Biochemistry*, 29 (32), pp.7387-401.
- Clubb, R.T., Thanabal, V. & Wagner, G. (1992) A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue ¹HN, ¹⁵N, and ¹³C' chemical shifts in ¹⁵N---¹³C-labelled proteins. *Journal of Magnetic Resonance*, 97, pp.213-217.
- Cornish-Bowden, A. (2002) Enthalpy-entropy compensation: a phantom phenomenon. *Journal of Biosciences*, 27 (2), pp.121-6.
- Corrêa, D.H.A. & Ramos, C.H.I. (2009) The use of circular dichroism spectroscopy to study protein folding, form and function. *African Journal of Biochemistry Research*, 3 (5), pp.164-173.
- Dembowski, N.J. & Kantrowitz, E.R. (1994) The use of alanine scanning mutagenesis to determine the role of the N-terminus of the regulatory chain in the heterotropic mechanism of Escherichia coli aspartate transcarbamoylase. *Protein Engineering*, 7 (5), pp.673-9.
- Doolittle, R.F. (1989) Redundancies in protein sequences. In: G. D. Fasman ed. *Prediction of protein structures and the principles of protein conformation*. Plenum Press, pp.599-623.
- D'Auria, S., Staiano, M., Varriale, A., Gonnelli, M., Marabotti, A., Rossi, M. & Strambini, G.B. (2008) The Tryptophan Phosphorescence of Porcine and Mutant Bovine Odorant-Binding Proteins: A Probe for the Local Protein Structure and Dynamics. *Journal of Proteome Research*, 7 (3), pp.1151-1158.
- Englander, S.W. & Wand, A.J. (1987) Main-chain-directed strategy for the assignment of ¹H NMR spectra of proteins. *Biochemistry*, 26 (19), pp.5953-8.
- Falconer, R.J., Penkova, A., Jelesarov, I. & Collins, B.M. (2010) Survey of the year 2008: applications of isothermal titration calorimetry. *Journal of Molecular Recognition*, 23 (5), pp.395-413.
- Favetta, P., Degoute, C.S., Perdrix, J.P., Dufresne, C., Boulieu, R. & Guitton, J. (2002) Propofol metabolites in man following propofol induction and maintenance. *British Journal of Anaesthesia*, 88 (5), pp.653-8.
- Felicioli, A., Ganni, M., Garibotti, M. & Pelosi, P. (1993) Multiple types and forms of odorant-binding proteins in the old-world porcupine *Hystrix cristata*. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry*, 105 (3-4), pp.775-784.

- Firestein, S. (2001) How the olfactory system makes sense of scents. *Nature*, 413, pp.211-218.
- Flower, D.R. (1996) The lipocalin protein family: structure and function. *Biochemical Journal*, 318, pp.1-14.
- Flower, D.R., North, A.C.. & Sansom, C.E. (2000) The lipocalin protein family: structural and sequence overview. *Biochim. Biophys. Acta*, 1482, pp.9-24.
- Forge, V., Hoshino, M., Kuwata, K., Arai, M., Kuwajima, K., Batt, C.A. & Goto, Y. (2000) Is folding of [beta]-lactoglobulin non-hierarchical? intermediate with native-like [beta]-sheet and non-native [alpha]-helix. *Journal of Molecular Biology*, 296 (4), pp.1039-1051.
- Freyer, M.W. & Lewis, E.A. (2008) Isothermal titration calorimetry: experimental design, data analysis, and probing macromolecule/ligand binding and kinetic interactions. *Methods in Cell Biology*, 84 (07), pp.79-113.
- Fuentes-Prior, P., Noeske-Jungblut, C., Donner, P., Schleuning, W.D., Huber, R. & Bode, W. (1997) Structure of the thrombin complex with triabin, a lipocalin-like exosite-binding inhibitor derived from a triatomine bug. *Proceedings of the National Academy of Sciences of the United States of America*, 94 (22), pp.11845-50.
- Godfrey, P.A., Malnic, B. & Buck, L.B. (2004) The mouse olfactory receptor gene family. *Proceedings of the National Academy of Sciences of the United States of America*, 101 (7), pp.2156-61.
- Golebiowski, J., Antonczak, S. & Cabrol-Bass, D. (2006) Molecular dynamics studies of odorant binding protein free of ligand and complexed to pyrazine and octenol. *Journal of Molecular Structure: THEOCHEM*, 763 (1-3), pp.165-174.
- Goto, N.K. & Kay, L.E. (2000) New developments in isotope labeling strategies for protein solution NMR spectroscopy. *Current Opinion in Structural Biology*, 10 (5), pp.585-92.
- Grasberger, B.L., Gronenborn, A.M. & Clore, G.M. (1993) Analysis of the Backbone Dynamics of Interleukin-8 by ¹⁵N Relaxation Measurements. *Journal of Molecular Biology*, 230 (2), pp.364-372.
- Grigorieff, N., Ceska, T.A., Downing, K.H., Baldwin, J.M. & Henderson, R. (1996) Electron-crystallographic Refinement of the Structure of Bacteriorhodopsin. *Journal of Molecular Biology*, 259 (3), pp.393-421.
- Grishaev, A., Tugarinov, V., Kay, L.E., Trewella, J. & Bax, A. (2008) Refined solution structure of the 82-kDa enzyme malate synthase G from joint NMR and synchrotron SAXS restraints. *Journal of Biomolecular NMR*, 40 (2), pp.95-106.

- Grolli, S., Merli, E., Conti, V., Scaltriti, E. & Ramoni, R. (2006) Odorant binding protein has the biochemical properties of a scavenger for 4-hydroxy-2-nonenal in mammalian nasal mucosa. *FEBS J*, 273, pp.5131-5142.
- Grzesiek, S. & Bax, A. (1992a) An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *Journal of Magnetic Resonance*, 99, pp.201-207.
- Grzesiek, S. & Bax, A. (1992b) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *Journal of the American Chemical Society*, 114 (16), pp.6291-6293.
- Grzesiek, S. & Bax, A. (1992c) Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *Journal of Magnetic Resonance*, 96, pp.432-440.
- Grzyb, J. & Latowski, D. (2006) Lipocalins – a family portrait. *Journal of Plant Physiology*, 163, pp.895-915.
- Guex, N. & Peitsch, M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, 18 (15), pp.2714-23.
- Gust, D., Moon, R.B. & Roberts, J.D. (1975) Applications of natural-abundance nitrogen-15 nuclear magnetic resonance to large biochemically important molecules. *Proceedings of the National Academy of Sciences of the United States of America*, 72 (12), pp.4696-700.
- Hajjar, E., Perahia, D., Nespoulous, C. & Robert, C.H. (2006) Odorant Binding and Conformational Dynamics in the Odorant-binding Protein. *Journal of Biological Chemistry*, 281 (40), pp.29929 -29937.
- Hamdani, E.. & Døving, K.B. (2007) The functional organization of the fish olfactory system. *Progress in Neurobiology*, 82 (2), pp.80-86.
- Hansson, B.S. (2002) A bug's smell - research into insect olfaction. *Trends in Neurosciences*, 25 (5), pp.270-4.
- Harris, C.L. (1981) Cysteine and growth inhibition of *Escherichia coli*: threonine deaminase as the target enzyme. *Journal of Bacteriology*, 145 (2), pp.1031-5.
- Henriques, D.A. & Ladbury, J.E. (2001) Inhibitors to the Src SH2 domain: a lesson in structure--thermodynamic correlation in drug design. *Archives of Biochemistry and Biophysics*, 390 (2), pp.158-68.
- Herr, F.M., Aronson, J. & Storch, J. (1996) Role of Portal Region Lysine Residues in Electrostatic Interactions between Heart Fatty Acid Binding Protein and Phospholipid Membranes. *Biochemistry*, 35 (4), pp.1296-1303.

- Hildebrand, J.G. & Shepherd, G.M. (1997) Mechanisms of olfactory discrimination: converging evidence for common principles across phyla. *Annual Review of Neuroscience*, 20 (1), pp.595-631.
- Homans, S.W. (2005) Probing the binding entropy of ligand-protein interactions by NMR. *European Journal of Chemical Biology*, 6 (9), pp.1585-91.
- Howard, H.H. (1993) Sec butyl acetate. In: *Handbook of environmental fate and exposure data for organic chemists*. pp.60-65.
- Janin, J. (1979) Surface and inside volumes in globular proteins. *Nature*, 277 (5696), pp.491-492.
- Jenkins, P.M., McEwen, D.P. & Martens, J.R. (2009) Olfactory cilia: linking sensory cilia function and human disease. *Chemical Senses*, 34 (5), pp.451-64.
- Jue, T., Mar, G.N.L.A., Han, K. & Yamamoto, Y. (1984) NMR study of the exchange rates of allosterically responsive labile protons in the heme pockets of haemoglobin A. *Biophysical Journal*, 46, pp.117-120.
- Karkhanis, V.A., Mascarenhas, A.. & Martinis, S.A. (2007) Amino acid toxicities of *Escherichia coli* that are prevented by leucyl-tRNA synthetase amino acid editing. *Journal of Bacteriology*, 189 (23), pp.8765-8.
- Katada, S., Hirokawa, T., Oka, Y., Suwa, M. & Touhara, K. (2005) Structural basis for a broad but selective ligand spectrum of a mouse olfactory receptor: mapping the odorant-binding site. *The Journal of Neuroscience*, 25 (7), pp.1806-15.
- Kato, A. & Touhara, K. (2009) Mammalian olfactory receptors: pharmacology, G protein coupling and desensitization. *Cellular and Molecular Life Sciences*, 66 (23), pp.3743-53.
- Kato, A., Katada, S. & Touhara, K. (2008) Amino acids involved in conformational dynamics and G protein coupling of an odorant receptor: targeting gain-of-function mutation. *Journal of Neurochemistry*, 107 (5), pp.1261-1270.
- Kaufman, K.M. & Sodetz, J.M. (1994) Genomic Structure of the Human Complement Protein C8.gamma.: Homology to the Lipocalin Gene Family. *Biochemistry*, 33 (17), pp.5162-5166.
- Kaupp, U.B. (2010) Olfactory signalling in vertebrates and insects: differences and commonalities. *Nature reviews: Neuroscience*, 11 (3), pp.188-200.
- Kay, L.E., Ikura, M., Tschudin, R. & Bax, A. (1990) Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *Journal of Magnetic Resonance*, 89, pp.496-514.
- Keil, T.A. (1999) Morphology and development of the peripheral olfactory organ. In: *Insect Olfaction*. pp.5-48.

- Kiefer, F., Arnold, K., Künzli, M., Bordoli, L. & Schwede, T. (2009) The SWISS-MODEL Repository and associated resources. *Nucleic Acids Research*, 37, pp.D387-92.
- Knopf, J., Gallagher, J. & Held, W.A (1983) Differential, multihormonal regulation of the mouse major urinary protein gene family in the liver. *Molecular and Cellular Biology*, 3 (12), pp.2232-40.
- Kouvatsos, N., Thurston, V., Ball, K., Oldham, N.J., Thomas, N.R. & Searle, M.S. (2007) Bile acid interactions with rabbit ileal lipid binding protein and an engineered helixless variant reveal novel ligand binding properties of a versatile beta-clam shell protein scaffold. *Journal of Molecular Biology*, 371 (5), pp.1365-77.
- Kramer, J., Lobkovsky, E.B. & Coates, G.W. (2006) Practical β -Lactone Synthesis: Epoxide Carbonylation at 1 atm. *Organic Letters*, 8 (17), pp.3709-3712.
- Krieger, J., Nickisch-Rosenegk, E., Mameli, M., Pelosi, P. & Breer, H. (1996) Binding proteins from the antennae of *Bombyx mori*. *Insect Biochemistry and Molecular Biology*, 26 (3), pp.297-307.
- Krishnamurthy, V.M., Bohall, B.R., Semetey, V. & Whitesides, G.M. (2006) The paradoxical thermodynamic basis for the interaction of ethylene glycol, glycine, and sarcosine chains with bovine carbonic anhydrase II: an unexpected manifestation of enthalpy/entropy compensation. *Journal of the American Chemical Society*, 128 (17), pp.5802-12.
- Krishnarajuna, B., Jaipuria, G., Thakur, A., D'Silva, P. & Atreya, H.S. (2011) Amino acid selective unlabeled for sequence specific resonance assignments in proteins. *Journal of Biomolecular NMR*, 49 (1), pp.39-51.
- Ku, T., Lu, P., Chan, C., Wang, T., Lai, S., Lyu, P. & Hsiao, N. (2009) Predicting melting temperature directly from protein sequences. *Computational Biology and Chemistry*, 33 (6), pp.445-50.
- Kuhn, N.J., Woodworth-Gutai, M., Gross, K.W. & Held, W.A. (1984) Subfamilies of the mouse major urinary protein (MUP) multi-gene family: sequence analysis of cDNA clones and differential regulation in the liver. *Nucleic Acids Research*, 12 (15), pp.6073-6090.
- Kuser, P.R., Franzoni, L., Ferrari, E., Spisni, A. & Polikarpov, I. (2001) The X-ray structure of a recombinant major urinary protein at 1.75 Å resolution. A comparative study of X-ray and NMR-derived structures. *Acta Crystallographica Section D Biological Crystallography*, 57 (12), pp.1863-1869.
- Kuwata, K., Li, H., Yamada, H., Batt, C.A., Goto, Y. & Akasaka, K. (2001) High pressure NMR reveals a variety of fluctuating conformers in [beta]-lactoglobulin. *Journal of Molecular Biology*, 305 (5), pp.1073-1083.

- Kyte, J. & Doolittle, R.F. (1982) A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, 157 (1), pp.105-132.
- Ladbury, J.E. (2004) Application of isothermal titration calorimetry in the biological sciences: things are heating up! *BioTechniques*, 37 (6), pp.885-7.
- Lartigue, A., Gruez, A., Briand, L., Blon, F., Bézirard, V., Walsh, M., Pernollet, J.C., Tegoni, Mariella & Cambillau, Christian (2004) Sulfur single-wavelength anomalous diffraction crystal structure of a pheromone-binding protein from the honeybee *Apis mellifera* L. *The Journal of Biological Chemistry*, 279 (6), pp.4459-64.
- Lefèvre, F., Rémy, M.H. & Masson, J.M. (1997) Alanine-stretch scanning mutagenesis: a simple and efficient method to probe protein structure and function. *Nucleic Acids Research*, 25 (2), pp.447-8.
- Lescop, E., Briand, L., Pernollet, J.C. & Guittet, E. (2009) Structural basis of the broad specificity of a general odorant-binding protein from honeybee. *Biochemistry*, 48 (11), pp.2431-41.
- Lewis, J.L. & Dahl, A.R. (2003) Olfactory Mucosa. In: R. L. Doty ed. *Handbook of olfaction and gustation*. pp.33-52.
- Li, L., Dantzer, J.J., Nowacki, J., O'Callaghan, B.J. & Meroueh, S.O. (2008) PDBcal: a comprehensive dataset for receptor-ligand interactions with three-dimensional structures and binding thermodynamics from isothermal titration calorimetry. *Chemical Biology & Drug Design*, 71 (6), pp.529-32.
- Lin, Y., Ahn, S., Murali, N., Brey, W., Bowers, C.R. & Warren, W.S. (2000) High-resolution, >1 GHz NMR in unstable magnetic fields. *Physical Review Letters*, 85 (17), pp.3732-5.
- Liou, H.L., Kahn, P.C. & Storch, J. (2002) Role of the helical domain in fatty acid transfer from adipocyte and heart fatty acid-binding proteins to membranes: analysis of chimeric proteins. *The Journal of Biological Chemistry*, 277 (3), pp.1806-15.
- Liu, T., Lin, Y., Wen, X., Jorissen, R.N. & Gilson, M.K. (2007) BindingDB: a web-accessible database of experimentally determined protein-ligand binding affinities. *Nucleic Acids Research*, 35 (Database issue), pp.D198-201.
- Löbel, D., Jacob, M., Volkner, M. & Breer, H. (2002) Odorants of different chemical classes interact with distinct odorant binding protein subtypes. *Chem. Senses*, 27, pp.39-44.
- Löbel, D., Marchese, S., Krieger, J., Pelosi, P. & Breer, H. (1998) Subtypes of odorant-binding proteins. *Agricultural Chemistry*, 324, pp.318-324.
- Löbel, D., Strotmann, J., Jacob, M. & Breer, H. (2001) Identification of a third rat odorant-binding protein (OBP3). *Chem. Senses*, 26, pp.673-680.

- Lücke, C., Franzoni, L., Abbate, F., Löhr, F., Ferrari, E., Sorbi, R.T., Rüterjans, H. & Spisni, A. (1999) Solution structure of a recombinant mouse major urinary protein. *European Journal of Biochemistry / FEBS*, 266 (3), pp.1210-1218.
- Ma, M., Grosmaître, X., Iwema, C.L., Baker, H., Greer, C. a & Shepherd, Gordon M (2003) Olfactory signal transduction in the mouse septal organ. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 23 (1), pp.317-24.
- Macek, P., Novák, P., Krízová, H., Zídek, L. & Sklenár, V. (2006) Molecular dynamics study of major urinary protein-pheromone interactions: a structural model for ligand-induced flexibility increase. *FEBS letters*, 580 (2), pp.682-4.
- Malnic, B., Godfrey, P.A. & Buck, L.B. (2004) The human olfactory receptor gene family. *Proceedings of the National Academy of Sciences of the United States of America*, 101 (8), pp.2584-9.
- Malnic, B., Hirono, J., Sato, T. & Buck, L.B. (1999) Combinatorial receptor codes for odors. *Cell*, 96 (5), pp.713-23.
- Marion, D., Driscoll, P.C., Kay, L.E., Wingfield, P.T., Bax, A., Gronenborn, A.M. & Clore, G.M. (1989) Overcoming the overlap problem in the assignment of ^1H NMR spectra of larger proteins by use of three-dimensional heteronuclear ^1H - ^{15}N Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy. *Biochemistry*, 28 (15), pp.6150-6.
- Marion, D., Kay, L.E., Sparks, S.W., Torchia, D.A. & Bax, A. (1989) Three-Dimensional Heteronuclear NMR of ^{15}N -Labeled Proteins. *Journal of the American Chemical Society*, (111), pp.1515-1517.
- Matarazzo, V., Zsürger, N., Guillemot, J.C., Clot-Faybesse, O., Botto, J.M., Dal Farra, C., Crowe, M., Demaille, J., Vincent, J.P., Mazella, J. & Ronin, C. (2002) Porcine odorant-binding protein selectively binds to a human olfactory receptor. *Chemical Senses*, 27 (8), pp.691-701.
- Meillour, P., Lagant, P., Cornard, J.P., Brimau, F., Le Danvic, C., Vergoten, G. & Michalski, J.C. (2009) Phenylalanine 35 and tyrosine 82 are involved in the uptake and release of ligand by porcine odorant-binding protein. *Biochimica et Biophysica Acta*, 1794 (8), pp.1142-50.
- Melo, F., Devos, D., Depiereux, E. & Feytmans, E. (1997) ANOLEA: a www server to assess protein structures. In: *Proceedings of International Conference on Intelligent Systems for Molecular Biology (American Association for Artificial Intelligence)*. pp.187-90.
- Mohanty, S., Zubkov, S. & Gronenborn, A.M. (2004) The solution NMR structure of *Antheraea polyphemus* PBP provides new insight into pheromone recognition by pheromone-binding proteins. *Journal of molecular biology*, 337 (2), pp.443-51.

- Mombaerts, P. (2004) Genes and ligands for odorant, vomeronasal and taste receptors. *Nature reviews: Neuroscience*, 5 (4), pp.263-78.
- Morrison, E.E. & Costanzo, R.M. (1990) Morphology of the human olfactory epithelium. *The Journal of Comparative Neurology*, 297 (1), pp.1-13.
- Muchmore, D.C., McIntosh, L.P., Russell, C.B., Anderson, E. & Dahlquist, F.W. (1989) Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance. *Methods in Enzymology*, 177, pp.44-75.
- Mucignat-Caretta, C. (2010) The rodent accessory olfactory system. *Journal Of Comparative Physiology A Neuroethology Sensory Neural And Behavioral Physiology*, 196 (10), pp.767-777.
- Muhandiram, D.R. & Kay, L.E. (1994) Gradient-Enhanced Triple-Resonance Three Dimensional NMR Experiments with Improved Sensitivity. *Journal of Magnetic Resonance*, 103, pp.203-216.
- Nespoulous, C., Briand, L. & Tran, V. (2004) Odorant Binding and Conformational Changes of a Rat Odorant-binding Protein. *Chemical Senses*, 29 (3), pp.189-198.
- Newcomer, M.E., Jones, T. a, Aqvist, J., Sundelin, J., Eriksson, U., Rask, L. & Peterson, P. a (1984) The three-dimensional structure of retinol-binding protein. *The EMBO journal*, 3 (7), pp.1451-4.
- Niimura, Y. & Nei, M. (2005) Evolutionary dynamics of olfactory receptor genes in fishes and tetrapods. *Proceedings of the National Academy of Sciences of the United States of America*, 102 (17), pp.6039-44.
- Novotny, M.V. & Wiesler, D. (1999) Positive identification of the puberty-accelerating pheromone of the house mouse : the volatile ligands associating with the major urinary protein. *Proceeding of the Royal Society London*, 266, pp.2017-2022.
- Ohloff, G. (1994) *Scent and fragrances: The fascination of odors and their chemical perspectives*. Springer-Verlag.
- Olsson, T.S.G., Williams, M.A., Pitt, W.R. & Ladbury, J.E. (2008) The thermodynamics of protein-ligand interaction and solvation: insights for ligand design. *Journal of Molecular Biology*, 384 (4), pp.1002-17.
- Pagano, A., Giannoni, P., Zambotti, A., Randazzo, N., Zerega, B., Cancedda, R. & Dozin, B. (2004) CALbeta, a novel lipocalin associated with chondrogenesis and inflammation. *European Journal of Cell Biology*, 81 (5), pp.264-272.
- Paolini, S., Scaloni, A., Amoresano, A., Marchese, S., Napolitano, E. & Pelosi, P. (1998) Amino acid sequence, post-translational modifications, binding and labelling of porcine odorant-binding protein. *Chemical Senses*, 23 (6), pp.689-98.

- Parisi, M., Mazzini, A., Tibor Sorbi, R., Ramoni, R., Grolli, S. & Favilla, R. (2005) Role of the disulphide bridge in folding, stability and function of porcine odorant binding protein: spectroscopic equilibrium studies on C63A/C155A double mutant. *Biochimica et Biophysica Acta*, 1750 (1), pp.30-9.
- Pelosi, P. (1994) Odour Binding Proteins: Structural aspects. In: *Annals New York Academy of Sciences*.
- Pelosi, P., Zhou, J., Ban, L. & Calvello, M. (2006) Soluble proteins in insect chemical communication. *Cellular and Molecular Life Sciences*, 63 (14), pp.1658-1676.
- Perez, M.D., Sanchez, L., Aranda, P., Ena, J.M., Oria, R. & Calvo, M. (1992) Effect of beta-lactoglobulin on the activity of pregastric lipase. A possible role for this protein in ruminant milk. *Biochimica et Biophysica Acta*, 24, pp.151-155.
- Perez-Miller, S., Zou, Q., Novotny, M.V. & Hurley, T.D. (2010) High resolution X-ray structures of mouse major urinary protein nasal isoform in complex with pheromones. *Protein Science*, 19 (8), pp.1469-79.
- Pervushin, K., Riek, R., Wider, G. & Wüthrich, K. (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proceedings of the National Academy of Sciences of the United States of America*, 94 (23), pp.12366-71.
- Pesenti, M., Spinelli, S., Bezirard, V., Briand, L., Pernollet, J., Tegoni, M. & Cambillau, C. (2008) Structural Basis of the Honey Bee PBP Pheromone and pH-induced Conformational Change. *Journal of Molecular Biology*, 380 (1), pp.158-169.
- Pesenti, M.E., Spinelli, S., Bezirard, V., Briand, L., Pernollet, J.C., Campanacci, V., Tegoni, M. & Cambillau, C. (2009) Queen bee pheromone binding protein pH-induced domain swapping favors pheromone release. *Journal of Molecular Biology*, 390 (5), pp.981-90.
- Van Petegem, F., Duderstadt, K.E., Clark, K., Wang, M. & Minor, D.L. (2008) Alanine-Scanning Mutagenesis Defines a Conserved Energetic Hotspot in the CaV α 1 AID-CaV β Interaction Site that Is Critical for Channel Modulation. *Structure*, 16 (2), pp.280-294.
- Phelan, M.M., McLean, L., Simpson, D.M., Hurst, J.L., Beynon, R. & Lian, L.Y. (2010) ¹H, ¹⁵N and ¹³C resonance assignment of darcin, a mouse major urinary protein. *Biomolecular NMR Assignments*, 4 (2), pp.13-15.
- Purves, D., Augustine, G.J., Fitzpatrick, D., Katz, L.C., LaMantia, A.S., McNamara, J.O. & Williams, M. (2001) *Neuroscience*.

- Rabe, S. (2004) Volatile Release from Liquids: A Comparison of In Vivo APCI-MS, In-mouth Headspace Trapping and In vitro Mouth Model Data. *Chemical Senses*, 29 (2), pp.163-173.
- Redondo, C., Vouropoulou, M., Evans, J. & Findlay, J.B.C. (2008) Identification of the retinol-binding protein (RBP) interaction site and functional state of RBPs for the membrane receptor. *The FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 22 (4), pp.1043-54.
- Renner, C., Schleicher, M., Moroder, L. & Holak, T. (2002) Practical aspects of the 2D 15N-{1H}-NOE experiment. *Journal of Biomolecular NMR*, 23 (1), pp.23-33.
- Rhodes, D.G., Bossio, R.E. & Laue, T.M. (2009) *Determination of size, molecular weight, and presence of subunits*. 1st ed. Elsevier Inc.
- Rose, G.D., Geselowitz, A.R., Lesser, G.J., Lee, R.H. & Zehfus, M.H. (1985) Hydrophobicity of amino acid residues in globular proteins. *Science*, 229 (4716), pp.834-838.
- Roy, J. & Laughton, C. (2010) Long-timescale molecular-dynamics simulations of the major urinary protein provide atomistic interpretations of the unusual thermodynamics of ligand binding. *Biophysical journal*, 99 (1), pp.218-26.
- Sambrook, J. & Russell, D. (2001) *Molecular cloning: A laboratory manual*.
- Sangamnatdej, S., Paesen, G.C., Slovak, M. & Nuttall, P.A. (2002) A high affinity serotonin- and histamine-binding lipocalin from tick saliva. *Insect Molecular Biology*, 11 (1), pp.79-86.
- Sato, K., Pellegrino, M., Nakagawa, T., Vossell, L.B. & Touhara, K. (2008) Insect olfactory receptors are heteromeric ligand-gated ion channels. *Nature*, 452 (7190), pp.1002-1006.
- Sattler, M. (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 34 (2), pp.93-158.
- Schmiedeberg, K., Shirokova, E., Weber, H.P., Schilling, B., Meyerhof, W. & Krautwurst, D. (2007) Structural determinants of odorant recognition by the human olfactory receptors OR1A1 and OR1A2. *Journal of Structural Biology*, 159 (3), pp.400-412.
- Schwede, T., Kopp, J., Guex, N. & Peitsch, M.C. (2003) SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Research*, 31 (13), pp.3381-3385.
- Shahan, K., Denaro, M., Gilmartin, M., Shi, Y. & Derman, E. (1987) Expression of six mouse major urinary protein genes in the mammary, parotid, sublingual,

submaxillary, and lachrymal glands and in the liver. *Molecular and Cellular Biology*, 7 (5), pp.1947-54.

Sharp, K. (2001) Entropy – enthalpy compensation : Fact or artifact ? *Protein Science*, 10, pp.661-667.

Sharrow, S.D., Novotny, M.V. & Stone, M.J. (2003) Thermodynamic analysis of binding between mouse major urinary protein-I and the pheromone 2-sec-butyl-4,5-dihydrothiazole. *Biochemistry*, 42 (20), pp.6302-9.

Soga, S., Shirai, H., Kobori, M. & Hirayama, N. (2007) Use of amino acid composition to predict ligand-binding sites. *Journal of Chemical Information and Modeling*, 47 (2), pp.400-6.

Spinelli, S., Ramoni, R., Grolli, S., Bonicel, J., Cambillau, C. & Tegoni, M. (1998) The structure of the monomeric porcine odorant binding protein sheds light on the domain swapping mechanism. *Biochemistry*, 37 (22), pp.7913-7918.

Staiano, M., D'Auria, S., Varriale, A., Rossi, M., Marabotti, A., Fini, C., Stepanenko, O.V., Kuznetsova, I.M. & Turoverov, K.K. (2007) Stability and dynamics of the porcine odorant-binding protein. *Biochemistry*, 46 (39), pp.11120-7.

Steinbrecht, R.A. (1998) Odorant-Binding Proteins: Expression and Function. *Annals of the New York Academy of Sciences*, 855 (1), pp.323-332.

Stöckmann, H., Bronowska, A., Syme, N.R., Thompson, G.S., Kalverda, A.P., Warriner, S.L. & Homans, S.W. (2008) Residual ligand entropy in the binding of p-substituted benzenesulfonamide ligands to bovine carbonic anhydrase II. *Journal of the American Chemical Society*, 130 (37), pp.12420-6.

Stüber, D., Matile, H. & Garotta, G. (1990) System for high-level production in *Escherichia coli* and rapid purification of recombinant proteins: application to epitope mapping, preparation of antibodies, and structure-function analysis. In: I. Lefkovits & B. Pernis eds. *Immunological Methods*. pp.121-152.

Tanio, M., Tanaka, R., Tanaka, T. & Kohno, T. (2009) Amino acid-selective isotope labeling of proteins for nuclear magnetic resonance study: proteins secreted by *Brevibacillus choshinensis*. *Analytical Biochemistry*, 386 (2), pp.156-60.

Taylor, A.J., Cook, D.J. & Scott, D.J. (2008) Role of Odorant Binding Proteins: Comparing Hypothetical Mechanisms with Experimental Data. *Chemosensory Perception*, 1 (2), pp.153-162.

Tcatchoff, L., Nespoulous, C. & Pernollet, J.C. (2006) A single lysyl residue defines the binding specificity of a human odorant-binding protein for aldehydes. *FEBS Letters*, 580, pp.2102-2108.

Tegoni, M., Pelosi, P., Vincent, F., Spinelli, S., Campanacci, V., Grolli, S., Ramoni, R. & Cambillau, C. (2000) Mammalian odorant binding proteins. *Biochim. Biophys. Acta*, 1482, pp.229-240.

- Tegoni, M., Ramoni, R., Bignetti, E., Spinelli, S. & Cambillau, C. (1996) Domain swapping creates a third putative combining site in bovine odorant binding protein dimer. *Nat Struct Mol Biol*, 3 (10), pp.863-867.
- Thode, A.B., Kruse, S.W., Nix, J.C. & Jones, D.N.M. (2008) The Role of Multiple Hydrogen-Bonding Groups in Specific Alcohol Binding Sites in Proteins: Insights from Structural Studies of LUSH. *Journal of Molecular Biology*, 376 (5), pp.1360-1376.
- Thomson, J.A. & Ladbury, J.E. (2004) Isothermal Titration Calorimetry: A Tutorial. In: *Biocalorimetry 2*. pp.37-57.
- Timm, D.E., Baker, L.J., Mueller, H., Zidek, L. & Novotny, M.V. (2001) Structural basis of pheromone binding to mouse major urinary protein (MUP-I). *Protein Science*, 10, pp.997-1004.
- Treweek, T.M., Rekas, A., Walker, M.J. & Carver, J.A. (2010) A quantitative NMR spectroscopic examination of the flexibility of the C-terminal extensions of the molecular chaperones, [alpha]A- and [alpha]B-crystallin. *Experimental Eye Research*, 91 (5), pp.691-699.
- Tsitsanou, K., Thireou, T., Drakou, C., Koussis, K., Keramioti, M., Leonidas, D., Eliopoulos, E., Iatrou, K. & Zographos, S. (2011) Anopheles gambiae odorant binding protein crystal complex with the synthetic repellent DEET: implications for structure-based design of novel mosquito repellents. *Cellular and Molecular Life Sciences*, pp.1-15.
- Tugarinov, V., Muhandiram, R., Ayed, A. & Kay, L.E. (2002) Four-dimensional NMR spectroscopy of a 723-residue protein: chemical shift assignments and secondary structure of malate synthase g. *Journal of the American Chemical Society*, 124 (34), pp.10025-35.
- Turin, L. (1996) A spectroscopic mechanism for primary olfactory reception. *Chemical Senses*, 21, pp.773-791.
- Turin, L. & Yoshii, F. (2003) Structure-odor relations: a modern perspective. In: *Handbook of olfaction and gustation*. pp.275-294.
- Ulland, S., Ian, E., Borg-Karlson, A.K. & Mustaparta, H. (2006) Discrimination between enantiomers of linalool by olfactory receptor neurons in the cabbage moth *Mamestra brassicae* (L.). *Chemical Senses*, 31 (4), pp.325-34.
- Urade, Y. & Eguchi, N. (2002) Lipocalin-type and hematopoietic prostaglandin D synthases as a novel example of functional convergence. *Prostaglandins and Other Lipid Mediators*, 68 (375-382).
- Vidic, J., Grosclaude, J., Persuy, M.A., Badonnel, K., Baly, C., Caillol, M., Salesse, R. & Pajot-augy, E. (2008) On a chip demonstration of a functional role for odorant binding protein in the preservation of olfactory receptor activity at high odorant concentration. *Lab on a Chip*, pp.678-688.

- Villar, H.O. & Kauvar, L.M. (1994) Amino acid preferences at protein binding sites. *FEBS letters*, 349 (1), pp.125-30.
- Vincent, F., Ramoni, R., Spinelli, S., Grolli, S., Tegoni, M., Cambillau, C. & Meillour, N.L. (2004) Crystal structures of bovine odorant-binding protein in complex with odorant molecules. *European Journal of Biochemistry*, 3842, pp.3832-3842.
- Vincent, F., Spinelli, S., Ramoni, R., Grolli, S., Pelosi, P., Cambillau, C. & Tegoni, M. (2000) Complexes of porcine odorant binding protein with odorant molecules belonging to different chemical classes. *Journal of Molecular Biology*, 300, pp.127-139.
- Vogt, R.G. & Riddiford, L.M. (1981) Pheromone binding and inactivation by moth antennae. *Nature*, 293 (5828), pp.161-163.
- Vogt, R.G., Callahan, F.E., Rogers, M.E. & Dickens, J.C. (1999) Odorant binding protein diversity and distribution among the insect orders, as indicated by LAP, an OBP-related protein of the true bug *Lygus lineolaris* (Hemiptera, Heteroptera). *Chemical Senses*, 24 (5), pp.481-95.
- Vosshall, L.B. & Stocker, R.F. (2007) Molecular architecture of smell and taste in *Drosophila*. *Annual Review of Neuroscience*, 30, pp.505-33.
- Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J. & Laue, E.D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins*, 59 (4), pp.687-96.
- Wang, L., Berne, B.J. & Friesner, R.A. (2011) Ligand binding to protein-binding pockets with wet and dry regions. *Proceedings of the National Academy of Sciences of the United States of America*, 108 (4), pp.1326-30.
- Waugh, D.S. (1996) Genetic tools for selective labeling of proteins with alpha-15N amino acids. *Journal of Biomolecular NMR*, 2, pp.184-192.
- Wetzel, C.H., Oles, M., Wellerdieck, C., Kuczkowiak, M., Gisselmann, G. & Hatt, H. (1999) Specificity and sensitivity of a human olfactory receptor functionally expressed in human embryonic kidney 293 cells and *Xenopus Laevis* oocytes. *The Journal of Neuroscience*, 19 (17), pp.7426-33.
- White, S., Briand, L., Scott, D.J. & Borysik, A.J. (2009) Structure of rat odorant-binding protein OBP1 at 1.6 Å resolution. *Acta Crystallographica. Section D, Biological Crystallography*, 65 (Pt 5), pp.403-10.
- Wicher, D., Schafer, R., Bauernfeind, R., Stensmyr, M.C., Heller, R., Heinemann, S.H. & Hansson, B.S. (2008) *Drosophila* odorant receptors are both ligand-gated and cyclic-nucleotide-activated cation channels. *Nature*, 452 (7190), pp.1007-1011.

- Wijmenga, S.S., Hallenga, K. & Hilbers, C.W. (1989) A three-dimensional heteronuclear multiple-quantum coherence homonuclear hartmann-hahn experiment. *Journal of Magnetic Resonance*, 84, pp.634-642.
- Wishart, D.S., Sykes, B.D. & Richards, F.M. (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *Journal of Molecular Biology*, 222, pp.311-333.
- Wittkind, M. & Mueller, L. (1993) HNCACB, a High Sensitivity 3D NMR Experiment to Correlate Amide-Proton and Nitrogen Resonances with the Alpha- and Beta-Carbon Resonances in Proteins. *Journal of Magnetic Resonance*, 101, pp.201-205.
- Wolfenden, R., Andersson, L., Cullis, P.M. & Southgate, C.C.B. (1981) Affinities of amino acid side chains for solvent water. *Biochemistry*, 20 (4), pp.849-855.
- Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*. Wiley Interscience.
- Wüthrich, K. & Wagner, G. (1982) Amide proton and surface conformation of the basic pancreatic trypsin inhibitor in solution. *Journal of Molecular Biology*, 160, pp.343-361.
- Wüthrich, K. & Wagner, G. (1979) Nuclear magnetic resonance of labile protons in the basic pancreatic trypsin inhibitor. *Journal of Molecular Biology*, 130 (1), pp.1-18.
- Yabuki, M., Portman, K.L., Scott, D.J., Briand, L. & Taylor, A.J. (2010) DyBOBS: A Dynamic Biomimetic Assay for Odorant-Binding to Odor-Binding Protein. *Chemosensory Perception*, 3 (2), pp.108-117.
- Yabuki, M., Scott, D.J., Briand, L. & Taylor, A.J. (2011) Dynamics of Odorant Binding to Thin Aqueous Films of Rat-OBP3. *Chemical Senses*, pp.1-13.
- Young, J.M. & Trask, B.J. (2002) The sense of smell: genomics of vertebrate odorant receptors. *Human Molecular Genetics*, 11 (10), pp.1153-60.
- Zanotti, G., Malpeli, G. & Berni, R. (1993) The Interaction of N-Ethyl Retinamide with Plasma Retinol-binding Crystal Structure of the Retinoid-RBP Complex Protein (RBP) and the at 1.9-A Resolution. *Journal of Biological Chemistry*, 268 (33), pp.24873-24879.
- Zarzo, M. (2007) The sense of smell: molecular basis of odorant recognition. *Biological Reviews of the Cambridge Philosophical Society*, 82 (3), pp.455-79.
- Zhang, X. & Firestein, S. (2002) The olfactory receptor gene superfamily of the mouse. *Nature: Neuroscience*, 5 (2), pp.124-133.
- Zhou, J.J., Robertson, G., He, X., Dufour, S., Hooper, A.M., Pickett, J.A., Keep, N.H. & Field, L.M. (2009) Characterisation of Bombyx mori Odorant-binding

proteins reveals that a general odorant-binding protein discriminates between sex pheromone components. *Journal of Molecular Biology*, 389 (3), pp.529-45.

Zhu, G., Xia, Y., Nicholson, L.K. & Sze, K.H. (2000) Protein Dynamics Measurements by TROSY-Based NMR Experiments. *Journal of Magnetic Resonance*, 143 (2), pp.423-426.

Zubkov, S., Gronenborn, A.M., Byeon, I.J.L. & Mohanty, S. (2005) Structural consequences of the pH-induced conformational switch in A.polyphemus pheromone-binding protein: mechanisms of ligand release. *Journal of Molecular Biology*, 354 (5), pp.1081-90.

Zuiderweg, E.R. & Fesik, S.W. (1989) Heteronuclear three-dimensional NMR spectroscopy of the inflammatory protein C5a. *Biochemistry*, 28, pp.2387-91.

Židek, L., Novotny, M.V. & Stone, M.J. (1999) Increased protein backbone conformational entropy upon hydrophobic ligand binding. *Nature Structural Biology*, 6 (12), pp.4-7.

Židek, L., Stone, M.J., Lato, S.M., Pagel, M.D., Miao, Z., Ellington, A.D. & Novotny, M.V. (1999) NMR Mapping of the Recombinant Mouse Major Urinary Protein I Binding Site Occupied by the Pheromone 2-sec-Butyl-4,5-dihydrothiazole. *Biochemistry*, 38 (31), pp.9850-9861.

Appendix I Primers used in this study

Primer name	Sequence
pQE31 sequencing forward	5' CCC GAA AAG TGC CAC CTG 3'
pQE31 sequencing reverse	5' TTC TGA GGT CAT TAC TGG 3'
F54A forward	5' GTC TTG GAG AAT TCC TTA GGC <u>GCC</u> ACG TTC CGT ATT AAG GAA AA 3'
F54A reverse	5' TTT TCC TTA ATA CGG AAC GTG GCG CCT AAG GAA TTC TCC AAG AC 3'
F56A forward	5' GAA TAT TT GTT GAG TAT GAC GGA GAG AAT ACA <u>GCG</u> ACT ATA CTT AAG AC 3'
F56A reverse	5' ACT CCA TTT TCC TTA ATA CGG GCC GTG AAG CCT AAG GAA TTC TC 3'
F90A forward	5' GAA TAT TTT GTT GAG TAT GAC GGA GAG AAT ACA <u>GCG</u> ACT ATA CTT AAG ACA GAC TAT GAC AAT T 3'
F90A reverse	5' AAT TGT CAT AGT CTG TCT TAA GTA TAG TCG CTG TAT TCT CTC CGT CAT ACT CAA CAA AAT ATT C 3'
F103A forward	5' G ACA GAC TAT GAC AAT TAT GTC ATG <u>GCG</u> CAT CTC GTT AAT GTC AAC AAC GGG G 3'
F103A reverse	5' C CCC GTT GTT GAC ATT AAC GAG ATG CGC CAT GAC ATA AT T GTC ATA GTC TGT C 3'
Y84A forward	5' GCA AAG GAT GGC GAA TAT TTT GTT GAG <u>GCG</u> GAC GGA GAG AAT ACA TTT ACT ATA C 3'
Y84A reverse	5' GTA TAG TAA ATG TAT TCT CTC CGT CCG CCT CAA CAA AAT ATT CGC CAT CCT TTG C 3'
Y84F forward	5' CA AAG GAT GGC GAA TAT TTT GTT GAG <u>TTT</u> GAC GGA GAG AAT AC 3'
Y84F reverse	5' GT ATT CTC TCC GTC AAA CTC AAC AAA ATA TTC GCC ATC CTT TG 3'
Y120A forward	5' CC TTC CAG CTG ATG GAG CTC <u>GCG</u> GGC AGA ACA AAG GAT CTG AG 3'
Y120A reverse	5' CT CAG ATC CTT TGT TCT GCC CGC GAG CTC CAT CAG CTG GAA GG 3'
M38L forward	5' TA GAA GAG AAC GGC AGC <u>CTG</u> AGA GTT TTT GTG CAG 3'
M38L reverse	5' CT GCA CAA AAA CTC TCA GGC TGC CGT TCT CTT CTA 3'
V40A forward	5' GAG AAC GGC AGC ATG AGA <u>GCG</u> TTT GTG CAG CAT ATC GAT G 3'
V40A reverse	5' CAT CGA TAT GCT GCA CAA ACG CTC TCA TGC TGC CGT TCT C 3'
L105N forward	5' ATG TCA TGT TTC ATC TCG <u>AAC</u> ATG TCA ACA ACG GGG AAA 3'
L105N reverse	5' TTT CCC CGT TGT TGA CAT GTT CGA GAT GAA ACA TGA CAT 3'
E118L forward	5' G GAA ACC TTC CAG CTG ATG <u>CTG</u> CTC TAT GGC AGA ACA AAG 3'
E118L reverse	5' C TTT GTT CTG CCA TAG AGC AGC ATC AGC TGG AAG GTT TCC 3'

Appendix II Ligands used in this study

Ligand	Sigma product code	Purity (%)	Methanol Used
2-isobutyl-3-methoxypyrazine	W313203	≥ 99	
2-isobutylthiazole	W313408	≥ 99	
4,5-dimethylthiazole	W327409	≥ 97	
Octan-1-ol	472328	≥ 99	< 1 %
Linalool	W263508	≥ 97	
Eugenol	W246719	≥ 98	< 1 %
Octanal	O5608	99	< 1 %
Ethylbutyrate	75563	≥ 99.5	
γ-hexalactone	303836	98	
γ-heptalactone	W253901	≥ 98	
γ-octalactone	W279617	≥ 97	
γ-nonolactone	W278106	≥ 98	
γ-decalactone	W236004	≥ 98	
γ-undecalactone	W309109	≥ 98	
γ-dodecalactone	W240001	≥ 97	< 1 %

Appendix III NMR acquisition parameters

Experiment	Points in Direct Dimension	Points in First Indirect Dimension	Points in Second Indirect Dimension	Number of scans	Spectral widths (ppm)	Spectrometer frequency	Notes
^1H watergate	2048 (^1H)	-	-	16	15 (^1H)	600 MHz	
$^1\text{H}/^{15}\text{N}$ TROSY	2048 (^1H)	128 (^{15}N)	-	32	13 (^1H), 40 (^{15}N)	600 MHz	
$^1\text{H}/^1\text{H}$ TOCSY	2048 (^1H)	512 (^1H)	-	64	13 (^1H), 13 (^1H)	600 MHz	Mix times: 60ms and 120ms
$^1\text{H}/^1\text{H}$ NOESY	2048 (^1H)	512 (^1H)	-	64	13 (^1H), 13 (^1H)	600 MHz	Mix times: 80ms and 120ms
$^1\text{H}/^{15}\text{N}$ TOCSY - TROSY	2048 (^1H)	32 (^{15}N)	128 (^1H)	16	13 (^1H), 40 (^{15}N)	600 MHz	Mix times: 60ms and 120ms
$^1\text{H}/^{15}\text{N}$ NOESY - TROSY	2048 (^1H)	32 (^{15}N)	128 (^1H)	16	13 (^1H), 40 (^{15}N)	600 MHz	Mix times: 80ms and 120ms
$^1\text{H}/^{13}\text{C}$ HSQC	2048 (^1H)	64 (^{13}C)	-	8	13 (^1H), 140 (^{13}C)	600 MHz	
HCC H TOCSY	2048 (^1H)	32 (^1H)	32 (^{13}C)	16	13 (^1H), 25 (^{13}C)	600 MHz	
HN(CA)CO	2048 (^1H)	40 (^{15}N)	96 (^{13}C)	64	13 (^1H), 40 (^{15}N), 20 (^{13}C)	600 MHz	

HNCO	2048 (¹ H)	40 (¹⁵ N)	96 (¹³ C)	16	13 (¹ H), 40 (¹⁵ N), 20 (¹³ C)	600 MHz	
CBCACONH	2048 (¹ H)	40 (¹⁵ N)	80 (¹³ C)	16	13 (¹ H), 40 (¹⁵ N), 80 (¹³ C)	600 MHz	
CBCANH	2048 (¹ H)	40 (¹⁵ N)	92 (¹³ C)	32	13 (¹ H), 40 (¹⁵ N), 80 (¹³ C)	600 MHz	Recorded on unbound OBP3 only
HNCA	2048 (¹ H)	36 (¹⁵ N)	96 (¹³ C)	16	13 (¹ H), 40 (¹⁵ N), 36 (¹³ C)	600 MHz	
HNCACB	2048 (¹ H)	36 (¹⁵ N)	98 (¹³ C)	64	13 (¹ H), 40 (¹⁵ N), 80 (¹³ C)	600 MHz	Recorded on bound OBP3 only
CACO	1024 (¹³ C)	256 (¹³ C)	-	16	32 (¹³ C), 32 (¹³ C)	500 MHz	Recorded on bound OBP3 only
C-CCCC TOCSY	1024 (¹³ C)	32 (¹³ C)	128 (¹³ C)	32	33 (¹³ C), 32 (¹³ C), 80 (¹³ C)	500 MHz	Recorded on bound OBP3 only. Mix times: 10ms and 20ms
HCACO	1024 (¹ H)	32 (¹³ C)	24 (¹³ C)	32	10 (¹ H), 32 (¹³ C), 32 (¹³ C)	500 MHz	Recorded on bound OBP3 only
TROSY NOE	2048 (¹ H)	128 (¹⁵ N)	-	16	13 (¹ H), 40 (¹⁵ N)	600 MHz	

Appendix IV Crystallisation screens

Commercial Crystal Screens (Qiagen):

- JCSG + series 1
- JCSG + series 2

Home screens (Scott White, Birmingham):

- 0.1 M imidazole, 1.6 M ammonium sulphate, 10 % glycerol, additives (pH 7)
- 0.1 M imidazole, 40 % polyethylene glycol (PEG) 400, 10 % glycerol, additives (pH 7)
- 0.1 M tris, 30% PEG 4000, 10 % glycerol, additives (pH 8)
- 0.1 M tris, 1.6 M ammonium sulphate, 10 % glycerol, additives (pH 8)
- 0.1 M imidazole, 30 % PEG 4000, 10 % glycerol, additives (pH 7)
- 0.1 M tris, 2 M sodium chloride, 10 % glycerol, additives (pH 8)
- 0.1 M tris, 40 % PEG 400, additives (pH 7)
- 0.1 M imidazole, 2 M NaCl, 10 % glycerol, additives (pH 7)

Appendix V Chemical shifts of unbound 1 mM OBP3 in 10 mM potassium phosphate, pH 6.1, 10 % D₂O, 0.04 % sodium azide (recorded at 298 K)

Residue	δH_N	δN_H	$\delta C\alpha$	$\delta C\beta$	δCO
Glu 1	8.39	121.07	53.98	27.60	173.66
Glu 2	8.02	121.90	53.88	27.87	173.11
Ala 3	8.27	126.71	49.20	17.56	174.33
Ser 4	8.10	113.97	55.15	62.89	171.54
Phe 5	8.14	123.24	56.92	36.69	174.41
Glu 6	8.43	121.73	54.97	27.08	174.09
Arg 7	7.86	120.47	54.00	27.60	174.30
Gly 8	8.01	109.12	42.98	-	171.34
Asn 9	8.09	118.43	49.69	35.93	172.07
Leu 10	7.97	123.04	52.91	40.23	173.15
Asp 11	8.35	127.34	49.47	37.17	173.55
Val 12	8.05	121.18	63.27	29.68	173.85
Asp 13	7.97	118.96	54.83	37.93	-
Lys 14	7.35	117.03	54.92	29.63	174.71
Leu 15	8.06	117.58	52.51	39.09	-
Asn 16	7.01	117.20	52.57	38.45	174.25
Gly 17	8.83	111.36	41.47	-	169.15
Asp 18	8.38	121.67	52.67	38.06	-
Trp 19	6.82	121.44	52.88	34.30	169.12
Phe 20	9.15	113.94	53.74	40.25	174.31
Ser 21	9.87	122.92	58.73	61.99	170.63
Ile 22	8.87	125.61	57.75	-	-
Val 23	7.63	114.48	57.63	34.12	171.47
Val 24	8.99	126.15	57.02	33.69	168.73
Ala 25	8.81	127.14	48.32	24.15	173.12
Ser 26	7.31	113.79	53.92	64.41	170.50
Asp 27	8.39	122.96	51.78	36.83	173.22
Lys 28	8.46	123.74	52.66	29.42	172.55
Arg 29	8.31	127.84	57.63	28.15	174.53
Glu 30	9.25	115.51	56.48	25.83	175.97
Lys 31	7.27	115.75	53.55	27.89	173.84
Ile 32	7.39	107.60	58.10	35.47	172.60
Glu 33	6.44	122.18	53.68	28.27	172.75
Glu 34	8.69	119.75	57.51	26.85	175.01
Asn 35	8.51	118.25	52.60	35.23	175.06
Gly 36	8.59	110.97	42.26	-	173.93
Ser 37	8.63	120.59	57.67	59.57	-
Met 38	7.90	115.31	50.47	28.27	172.78
Arg 39	7.13	127.10	51.17	25.37	170.04
Val 40	6.60	110.85	55.31	31.10	172.55
Phe 41	-	-	-	-	-

Val 42	8.85	131.46	60.49	-	-
Gln 43	-	-	-	-	-
His 44	-	-	-	-	-
Ile 45	-	-	-	-	-
Asp 46	9.05	127.51	49.98	41.18	172.00
Val 47	8.61	125.62	60.97	28.83	172.47
Leu 48	7.77	131.63	51.12	40.43	173.70
Glu 49	8.29	120.92	56.98	26.50	174.64
Asn 50	8.71	117.79	50.64	37.28	171.57
Ser 51	7.33	112.11	55.41	64.00	169.83
Leu 52	8.92	122.67	51.18	43.75	171.87
Gly 53	9.49	112.82	41.99	-	169.95
Phe 54	9.07	123.21	53.98	41.57	173.04
Thr 55	7.57	118.03	59.58	67.97	-
Phe 56	8.86	126.23	53.48	41.94	171.91
Arg 57	8.64	115.98	52.21	32.32	171.26
Ile 58	8.90	120.88	57.66	40.97	169.05
Lys 59	8.92	129.28	52.92	30.78	172.66
Glu 60	8.48	127.51	52.26	29.42	173.53
Asn 61	9.50	126.80	51.62	34.88	172.77
Gly 62	8.46	102.70	42.74	-	170.94
Val 63	7.60	121.39	58.61	32.29	173.18
Cys 64	8.92	127.20	55.65	42.04	171.72
Thr 65	9.28	121.56	58.78	68.93	169.34
Glu 66	8.50	125.89	52.24	31.08	172.31
Phe 67	9.19	124.49	53.76	38.39	169.08
Ser 68	8.56	115.20	53.98	63.71	170.72
Leu 69	8.73	122.11	51.02	45.12	172.96
Val 70	8.90	124.73	59.74	30.52	171.10
Ala 71	9.41	130.96	47.05	20.13	173.49
Asp 72	8.78	123.89	51.62	41.10	173.45
Lys 73	8.14	122.96	54.78	30.48	174.94
Thr 74	7.86	116.45	58.47	67.07	171.29
Ala 75	8.03	119.83	50.56	16.27	175.08
Lys 76	7.75	120.74	52.30	30.77	172.69
Asp 77	8.26	124.77	53.32	37.93	174.42
Gly 78	8.49	112.30	43.24	-	168.73
Glu 79	7.35	120.94	52.51	29.76	171.82
Tyr 80	-	-	-	-	-
Phe 81	-	-	-	-	-
Val 82	8.30	124.73	57.84	32.64	169.50
Glu 83	8.82	132.12	53.30	26.93	173.50
Tyr 84	8.68	127.05	57.19	35.86	170.19
Asp 85	8.60	130.59	51.79	36.55	-
Gly 86	7.85	104.51	39.78	-	170.15

Glu 87	8.29	118.50	53.48	28.83	172.80
Asn 88	7.48	120.81	-	-	-
Thr 89	8.57	116.75	58.28	68.22	-
Phe 90	9.10	121.35	-	-	-
Thr 91	9.37	111.76	-	-	-
Ile 92	8.63	123.02	-	-	-
Leu 93	-	-	-	-	-
Lys 94	6.84	114.21	52.75	34.78	171.64
Thr 95	8.35	117.73	57.74	67.65	-
Asp 96	6.58	124.28	49.94	39.78	176.28
Tyr 97	8.71	117.79	-	-	-
Asp 98	9.18	120.88	52.56	39.86	173.15
Asn 99	9.02	117.14	53.83	39.26	171.50
Tyr 100	9.25	117.34	54.11	41.92	171.66
Val 101	8.50	120.35	58.52	31.37	168.08
Met 102	-	-	-	-	-
Phe 103	-	-	-	-	-
His 104	-	-	-	-	-
Leu 105	-	-	-	-	-
Val 106	9.34	124.79	-	-	-
Asn 107	8.23	119.44	52.11	37.19	174.01
Val 108	6.73	113.44	52.19	28.09	168.95
Asn 109	8.47	124.50	50.06	37.77	173.06
Asn 110	9.19	126.53	51.56	34.55	172.65
Gly 111	8.40	104.46	42.78	-	170.99
Glu 112	7.67	121.47	52.13	29.47	173.31
Thr 113	8.67	118.97	57.24	69.26	170.71
Phe 114	8.11	120.85	54.55	38.32	169.37
Gln 115	10.16	118.24	51.21	32.16	172.32
Leu 116	9.44	127.06	50.96	45.35	-
Met 117	-	-	-	-	-
Glu 118	-	-	-	-	-
Leu 119	-	-	-	-	-
Tyr 120	8.68	126.04	49.23	36.79	173.11
Gly 121	9.65	107.28	41.17	-	171.75
Arg 122	8.67	120.96	53.40	28.67	174.02
Thr 123	7.42	107.76	56.19	68.64	169.97
Lys 124	7.57	114.98	56.11	29.62	172.33
Asp 125	7.51	117.37	49.07	41.11	171.70
Leu 126	8.71	120.89	49.93	46.38	172.88
Ser 127	8.12	115.44	55.28	61.95	172.50
Ser 128	8.91	119.01	59.07	-	172.51
Asp 129	8.28	121.57	54.66	37.86	-
Ile 130	7.38	122.35	61.40	34.51	175.26
Lys 131	7.56	119.97	58.70	29.22	176.54

Glu 132	8.24	122.00	55.50	25.56	175.42
Lys 133	7.74	122.73	57.30	29.65	177.25
Phe 134	8.15	119.47	58.46	36.83	174.04
Ala 135	8.48	125.09	53.47	15.35	176.91
Lys 136	8.15	116.89	57.73	29.97	-
Leu 137	7.36	122.27	54.89	39.28	177.34
Cys 138	8.04	123.19	61.29	22.95	174.85
Val 139	7.95	123.52	63.03	29.27	-
Ala 140	7.76	124.03	51.74	15.36	176.01
His 141	7.48	115.60	54.69	25.65	171.30
Gly 142	7.85	107.79	42.94	-	169.89
Ile 143	7.82	124.31	58.06	36.11	171.34
Thr 144	7.84	116.57	58.51	67.01	173.62
Arg 145	8.49	120.80	56.55	27.28	175.47
Asp 146	8.14	117.00	52.99	37.02	-
Asn 147	7.89	121.37	49.65	35.17	169.04
Ile 148	6.75	118.34	59.06	35.69	173.36
Ile 148	9.44	129.96	59.06	35.69	173.36
Asp 150	8.63	127.77	51.26	37.35	174.76
Leu 151	8.94	128.74	52.74	40.17	176.40
Thr 152	8.98	113.00	62.91	67.31	172.95
Lys 153	7.21	119.43	52.48	29.20	173.10
Thr 154	7.34	112.71	56.82	68.95	170.16
Asp 155	7.61	115.50	-	-	-
Arg 156	-	-	-	-	-
Cys 157	-	-	-	-	-
Leu 158	7.88	117.68	53.21	39.55	175.30
Gln 159	8.35	121.23	53.05	26.72	171.78
Ala 160	7.80	132.24	51.05	17.36	179.53

Appendix VI Chemical shifts of 2-isobutylthiazole bound OBP3 (2 mM to 1 mM) in 10 mM potassium phosphate, pH 6.1, 10 % D₂O, 0.04 % sodium azide (recorded at 298 K)

Residue	δH_{α}	δH_N	δN_H	δC_{α}	δC_{β}	δCO
Glu 1	4.26	8.39	121.04	53.93	27.50	173.74
Glu 2	4.17	8.04	121.74	53.91	28.12	173.18
Ala 3	4.38	8.27	126.39	49.23	17.45	174.39
Ser 4	5.16	8.09	113.88	55.31	62.38	171.70
Phe 5	4.47	8.14	123.14	56.77	36.45	174.29
Glu 6	4.06	8.45	121.72	54.95	27.12	174.07
Arg 7	4.16	7.88	120.38	53.97	27.88	174.38
Gly 8	3.88	8.03	109.08	42.98	-	171.36
Asn 9	4.72	8.11	118.47	49.72	35.91	172.05
Leu 10	4.12	7.98	123.03	52.78	40.28	173.22
Asp 11	4.73	8.37	127.18	49.48	37.32	174.26
Val 12	3.54	8.06	121.14	63.20	29.07	173.82
Asp 13	4.19	7.97	118.79	54.81	37.76	176.05
Lys 14	4.00	7.36	116.88	54.86	29.35	174.72
Leu 15	4.56	8.06	117.55	52.48	39.27	172.99
Asn 16	4.28	7.00	116.96	52.55	38.21	174.22
Gly 17	-	8.90	111.51	41.51	-	169.26
Asp 18	-	8.37	121.75	52.65	38.13	173.09
Trp 19	4.56	6.83	121.37	52.84	34.10	169.09
Phe 20	4.16	9.16	113.85	53.65	40.28	174.41
Ser 21	-	9.90	122.75	58.95	61.35	-
Ile 22	4.31	8.91	125.49	57.73	31.86	172.18
Val 23	5.09	7.58	115.21	57.61	33.19	171.70
Val 24	5.35	9.08	126.96	57.19	33.66	168.67
Ala 25	5.50	8.87	127.24	48.41	23.81	173.05
Ser 26	5.08	7.32	113.78	53.99	64.08	170.52
Asp 27	4.68	8.38	122.61	51.58	36.83	172.98
Lys 28	-	-	-	-	-	-
Arg 29	3.28	8.33	127.07	57.62	27.77	174.90
Glu 30	3.93	9.24	115.34	56.45	25.77	176.01
Lys 31	3.87	7.28	115.73	53.49	27.71	173.80
Ile 32	4.23	7.39	107.32	58.09	35.51	172.59
Glu 33	4.01	6.44	122.02	53.58	28.39	172.77
Glu 34	3.73	8.68	119.72	57.50	26.65	175.07
Asn 35	4.49	8.53	118.36	52.70	35.04	172.41
Gly 36	3.50,4.07	8.65	111.20	42.27	-	173.93
Ser 37	-	8.58	120.22	57.70	59.60	172.27
Met 38	-	7.91	115.72	50.89	28.35	172.14
Arg 39	4.00	6.96	125.28	50.97	25.00	170.39

Val 40	4.09	6.77	110.55	55.48	30.86	172.31
Phe 41	5.22	8.67	120.82	50.45	35.95	174.98
Val 42	3.64	9.17	131.62	60.41	29.06	169.12
Gln 43	4.53	8.23	117.70	50.94	28.70	173.86
His 44	4.97	6.66	113.11	52.20	28.15	168.97
Ile 45	4.31	8.45	124.36	58.02	37.82	171.00
Asp 46	-	9.10	127.48	49.74	41.17	172.03
Val 47	3.97	8.63	125.60	61.18	28.68	172.68
Leu 48	4.69	7.71	131.55	50.78	40.45	173.60
Glu 49	3.91	8.27	120.94	56.82	26.39	174.58
Asn 50	4.80	8.78	118.70	50.87	37.08	171.45
Ser 51	5.45	7.34	112.11	55.43	63.92	170.01
Leu 52	4.80	8.93	122.28	51.10	44.52	172.02
Gly 53	3.57,4.99	9.52	112.62	41.80	-	169.71
Phe 54	4.67	9.07	121.97	53.83	42.91	172.93
Thr 55	4.56	7.50	117.46	59.42	67.52	169.65
Phe 56	4.84	8.88	125.60	54.14	42.44	171.62
Arg 57	5.74	8.69	115.92	52.16	32.05	171.06
Ile 58	4.40	8.87	120.67	57.84	40.23	169.17
Lys 59	4.91	8.97	129.14	52.92	30.75	172.70
Glu 60	4.56	8.47	127.33	52.15	29.01	173.58
Asn 61	4.29	9.51	126.82	51.59	34.85	172.71
Gly 62	3.45,4.16	8.45	102.58	42.78	-	170.99
Val 63	4.26	7.60	121.36	58.70	32.06	173.27
Cys 64	5.11	8.95	127.25	55.76	41.99	171.77
Thr 65	4.57	9.29	121.40	58.77	68.77	169.30
Glu 66	5.54	8.46	125.36	52.18	31.22	172.30
Phe 67	5.04	9.18	124.13	53.78	38.63	169.31
Ser 68	5.53	8.62	115.25	54.14	63.57	170.69
Leu 69	4.96	8.84	121.49	51.05	45.56	173.37
Val 70	4.74	8.93	123.99	59.96	30.81	171.28
Ala 71	5.18	9.56	130.78	47.09	20.87	173.57
Asp 72	-	8.78	123.43	51.37	41.39	173.50
Lys 73	3.84	8.13	122.88	54.78	30.38	175.00
Thr 74	4.55	7.86	116.36	57.79	66.78	171.34
Ala 75	4.12	8.06	120.16	50.49	15.96	175.12
Lys 76	4.21	7.74	120.66	52.40	30.37	172.87
Asp 77	-	8.29	125.00	49.89	38.12	174.44
Gly 78	3.63,4.06	8.46	112.25	42.97	-	169.24
Glu 79	5.00	7.40	120.93	52.72	29.41	172.04
Tyr 80	-	-	-	-	-	-
Phe 81	-	-	-	53.27	41.24	172.18
Val 82	4.39	8.29	124.55	57.71	32.72	169.42
Glu 83	4.53	8.82	132.20	53.47	27.22	173.54
Tyr 84	4.03	8.53	126.98	57.52	35.69	170.27

Asp 85	3.70	8.52	130.58	51.66	36.89	172.31
Gly 86	3.57,4.38	7.81	104.04	39.92	-	170.21
Glu 87	4.12	8.29	118.85	53.40	29.15	172.65
Asn 88	5.93	7.42	120.67	49.56	39.25	171.37
Thr 89	5.59	8.52	115.74	57.45	68.12	171.89
Phe 90	6.06	9.13	121.37	53.00	40.75	169.60
Thr 91	5.16	9.32	110.65	56.77	69.96	171.20
Ile 92	-	8.66	123.40	58.98	-	173.36
Leu 93	4.16			54.03	-	174.60
Lys 94	4.21	6.81	113.96	52.79	34.73	171.88
Thr 95	4.11	8.35	117.82	57.75	67.31	166.65
Asp 96	-	6.59	124.29	49.85	39.55	176.22
Tyr 97	3.67	8.67	117.62	62.18	34.86	173.97
Asp 98	-	9.18	120.82	52.59	39.98	#N/A
Asn 99	5.04	9.02	117.00	53.88	39.30	171.53
Tyr 100	-	9.26	117.21	54.17	41.73	172.23
Val 101	4.33	8.51	120.28	58.53	31.27	168.08
Met 102	-	-	-	-	-	-
Phe 103	-	-	-	-	-	-
His 104	-	-	-	-	-	-
Leu 105	-	-	-	-	-	-
Val 106		9.39	124.81	-	-	-
Asn 107	-	8.29	119.37	-	-	-
Val 108	-	6.66	113.11	-	-	-
Asn 109	4.76	8.45	124.36	50.15	37.62	173.15
Asn 110	4.23	9.18	126.36	51.61	34.57	172.55
Gly 111	3.48,4.03	8.39	104.38	42.60	-	171.05
Glu 112	4.65	7.68	121.37	52.19	29.41	173.39
Thr 113	5.80	8.67	118.59	56.90	69.01	170.62
Phe 114	5.01	8.07	120.03	54.67	38.65	169.45
Gln 115	5.24	10.29	118.16	51.42	32.72	172.15
Leu 116	5.25	9.48	127.08	50.56	45.84	172.11
Met 117	5.36	9.39	124.74	51.90	34.27	172.76
Glu 118	5.04	8.29	119.37	52.11	34.20	169.71
Leu 119	4.66	-	-	50.30	40.29	171.38
Tyr 120	6.14	8.68	126.70	49.24	36.64	173.34
Gly 121	3.73	9.70	107.37	-	-	-
Arg 122	3.85	8.67	120.95	53.40	28.92	174.05
Thr 123	4.35	7.41	107.67	56.14	68.38	170.05
Lys 124	-	7.58	115.09	56.13	29.73	172.41
Asp 125	4.84	7.51	117.24	49.05	40.92	171.77
Leu 126	4.78	8.72	120.81	49.87	46.32	172.96
Ser 127	4.18	8.12	115.37	55.29	61.56	172.56
Ser 128	-	8.92	118.95	59.20	-	-
Asp 129	4.19	8.27	121.65	54.67	37.88	176.00

Ile 130	3.43	7.37	122.14	61.35	34.24	175.31
Lys 131	3.64	7.56	119.86	58.65	29.20	176.55
Glu 132	4.44	8.26	121.85	55.45	25.43	175.44
Lys 133	3.82	7.75	122.57	57.30	29.37	177.23
Phe 134	-	8.14	119.19	-	-	-
Ala 135	3.75	8.45	125.13	53.46	15.13	176.93
Lys 136	3.71	8.16	116.83	57.78	29.62	177.55
Leu 137	4.07	7.36	121.97	54.87	39.00	177.25
Cys 138	4.06	8.02	123.23	60.95	22.98	175.01
Val 139	3.85	7.97	123.60	63.03	29.10	178.33
Ala 140	4.05	7.77	123.95	51.70	15.17	176.02
His 141	4.34	7.49	115.37	54.67	25.78	171.32
Gly 142	3.63,4.17	7.87	107.79	43.00	-	171.52
Ile 143	4.00	7.82	124.11	58.10	35.76	171.26
Thr 144	4.11	7.86	116.64	58.50	66.80	173.61
Arg 145	3.77	8.49	120.82	56.59	26.71	175.48
Asp 146	4.40	8.14	117.47	52.83	36.99	173.23
Asn 147	4.55	7.91	121.31	49.55	35.03	169.02
Ile 148	4.28	6.76	118.20	59.09	36.32	173.37
Ile 148	4.28	9.45	129.82	59.09	36.32	173.37
Asp 150	4.80	8.65	127.79	51.20	37.32	174.76
Leu 151	4.23	8.93	128.73	52.74	39.89	176.45
Thr 152	3.79	8.99	112.91	62.85	66.48	173.05
Lys 153	4.41	7.20	119.30	52.47	29.25	173.14
Thr 154	4.10	7.35	112.45	56.95	68.34	170.18
Asp 155	-	7.59	114.95	-	-	-
Arg 156	-	-	-	-	-	-
Cys 157	4.15	-	-	53.70	-	172.36
Leu 158	4.09	7.88	117.65	53.18	39.61	175.25
Gln 159	4.22	8.34	121.08	53.12	26.70	171.84
Ala 160	3.97	7.78	132.07	51.09	17.33	179.55

Appendix VII Chemical shifts perturbations of NH resonances of OBP3 upon ligand binding

Residue	CSP IBT	CSP γ -hept	CSP γ -oct	CSP γ -non	CSP γ -dec	CSP γ -undec
Glu 1	0.01	0.02	0.02	0.02	0.02	0.02
Glu 2	0.03	0.01	0.01	0.01	0.01	0.03
Ala 3	0.05	0.02	0.02	0.02	0.01	0.01
Ser 4	0.02	0.02	0.02	0.01	0.01	0.00
Phe 5	0.02	0.03	0.04	0.02	0.03	0.03
Glu 6	0.02	-	-	-	-	-
Arg 7	0.02	0.03	0.05	0.05	0.04	0.02
Gly 8	0.02	0.02	0.03	0.04	0.03	0.01
Asn 9	0.02	0.01	0.02	0.03	0.02	0.01
Leu 10	0.01	0.05	0.01	0.02	0.01	0.03
Asp 11	0.03	0.03	0.05	0.06	0.01	0.02
Val 12	0.01	0.01	0.02	0.01	0.02	-
Asp 13	0.03	0.01	0.01	0.02	0.01	0.01
Lys 14	0.02	0.00	0.01	0.02	0.02	0.02
Leu 15	0.01	0.01	0.01	0.02	0.02	0.01
Asn 16	0.04	0.01	0.01	0.04	0.04	0.01
Gly 17	0.07	0.01	0.03	0.06	0.07	0.07
Asp 18	0.02	0.04	0.04	0.05	0.05	0.06
Trp 19	0.02	0.02	0.02	0.04	0.04	0.03
Phe 20	0.02	0.02	0.03	0.04	0.04	0.02
Ser 21	0.04	0.06	0.07	0.08	0.07	0.07
Ile 22	0.04	0.01	0.05	0.07	0.07	0.10
Val 23	-	-	-	-	-	-
Val 24	0.16	0.02	0.04	0.03	0.04	-
Ala 25	0.07	0.01	0.01	0.02	0.04	0.07
Ser 26	0.01	0.10	0.09	0.08	0.05	0.03
Asp 27	0.05	0.05	0.06	0.08	0.08	-
Lys 28	-	-	-	-	-	-
Arg 29	0.12	0.10	0.10	0.10	0.13	0.13
Glu 30	0.03	0.03	0.04	0.01	0.02	0.03
Lys 31	0.01	0.01	0.02	0.02	0.01	0.04
Ile 32	0.04	0.05	0.05	0.05	0.07	0.09
Glu 33	0.02	0.02	0.01	0.03	0.03	0.02
Glu 34	0.02	0.03	0.01	0.03	0.03	-
Asn 35	0.02	0.03	0.02	0.05	0.05	0.06
Gly 36	0.07	0.04	0.03	0.03	-	0.04
Ser 37	0.08	0.09	0.11	0.12	0.14	0.16
Met 38	0.07	0.09	0.05	0.04	0.03	0.06
Arg 39	0.33	0.18	0.21	0.19	0.17	0.09
Val 40	0.18	0.10	0.05	0.03	0.10	0.23
Phe 41	-	-	-	-	-	-

Val 42	0.32	0.22	0.20	0.17	0.12	0.02
Gln 43	-	-	-	-	-	-
His 44	-	-	-	-	-	-
Ile 45	-	-	-	-	-	-
Asp 46	0.05	0.02	0.03	0.04	0.05	0.04
Val 47	0.02	0.01	0.02	0.03	0.05	0.04
Leu 48	0.06	0.04	0.05	0.05	0.06	0.07
Glu 49	0.02	0.02	0.03	0.02	0.02	0.03
Asn 50	0.16	0.04	0.06	0.12	0.13	0.12
Ser 51	0.02	0.02	0.03	0.03	0.03	0.02
Leu 52	0.06	0.01	0.02	0.03	0.04	0.06
Gly 53	0.04	0.08	0.07	0.07	0.07	0.10
Phe 54	0.19	0.10	0.15	0.18	0.23	-
Thr 55	0.11	0.02	0.07	0.10	0.10	0.11
Phe 56	0.10	0.10	0.11	0.12	0.11	0.09
Arg 57	0.05	0.07	0.08	0.10	0.07	0.06
Ile 58	0.04	0.01	0.02	0.05	0.08	0.08
Lys 59	0.05	0.04	0.02	0.03	0.02	0.06
Glu 60	0.03	0.01	0.03	0.03	0.04	0.05
Asn 61	0.01	0.02	0.01	0.00	0.01	0.02
Gly 62	0.02	0.02	0.01	0.01	0.01	0.01
Val 63	0.00	0.02	0.01	0.01	0.02	0.03
Cys 64	0.03	0.02	0.02	0.03	0.02	0.01
Thr 65	0.03	0.02	0.01	0.02	0.00	0.03
Glu 66	0.09	0.05	0.05	0.02	0.01	0.08
Phe 67	0.06	0.03	0.05	0.05	0.08	0.13
Ser 68	0.06	0.01	0.05	0.06	0.06	0.04
Leu 69	0.15	0.08	0.07	0.09	0.11	0.15
Val 70	0.12	0.05	0.08	0.16	0.18	0.23
Ala 71	0.15	0.02	0.04	0.08	0.10	-
Asp 72	0.07	0.03	0.04	0.04	0.06	0.09
Lys 73	0.01	-	-	-	-	-
Thr 74	0.02	0.01	0.01	0.02	-	-
Ala 75	0.06	-	-	-	-	-
Lys 76	0.01	0.01	0.01	0.02	0.01	0.01
Asp 77	0.05	0.03	0.01	0.01	0.01	0.04
Gly 78	0.03	0.02	0.03	0.01	0.00	0.04
Glu 79	0.05	0.02	0.01	0.02	0.01	0.01
Tyr 80	-	-	-	-	-	-
Phe 81	-	-	-	-	-	-
Val 82	0.03	0.03	0.04	0.09	0.03	0.03
Glu 83	0.01	0.03	0.04	0.03	0.04	0.06
Tyr 84	0.14	0.15	-	-	-	-
Asp 85	0.09	0.10	0.10	0.09	0.09	0.08

Gly 86	0.08	0.03	0.03	0.02	0.01	0.03
Glu 87	0.05	0.12	0.09	0.03	0.01	0.05
Asn 88	0.07	0.08	0.09	0.06	0.09	0.11
Thr 89	0.16	0.16	0.15	0.04	-	0.03
Phe 90	0.03	0.30	0.32	0.40	0.41	0.36
Thr 91	0.18	0.10	0.11	0.09	-	0.11
Ile 92	0.06	0.14	0.12	0.13	0.15	-
Leu 93	-	-	-	-	-	-
Lys 94	0.05	0.01	0.02	0.03	0.03	0.02
Thr 95	0.01	0.01	0.02	0.03	0.05	0.04
Asp 96	0.01	0.02	0.02	0.02	0.03	0.01
Tyr 97	0.05	0.12	0.13	0.09	0.02	0.13
Asp 98	0.01	0.01	0.01	0.01	0.01	0.01
Asn 99	0.02	0.01	0.01	0.02	0.02	0.02
Tyr 100	0.02	0.02	0.02	0.02	0.02	0.02
Val 101	0.01	-	-	-	-	-
Met 102	-	-	-	-	-	-
Phe 103	-	-	-	-	-	-
His 104	-	-	-	-	-	-
Leu 105	-	-	-	-	-	-
Val 106	0.05	0.19	0.17	0.18	0.18	0.21
Asn 107	0.07	0.09	0.12	0.12	0.15	0.21
Val 108	0.09	0.03	0.02	0.03	0.03	0.07
Asn 109	0.03	0.02	0.02	0.03	0.05	0.07
Asn 110	0.03	0.02	0.03	0.03	0.02	0.02
Gly 111	0.01	0.01	0.01	0.01	0.01	0.01
Glu 112	0.01	0.02	0.01	0.03	0.03	0.01
Thr 113	0.06	0.13	0.14	0.11	0.00	0.05
Phe 114	0.13	0.19	0.18	0.17	0.17	0.17
Gln 115	0.13	0.17	0.18	0.17	0.14	0.03
Leu 116	0.04	0.26	0.24	0.14	0.07	0.08
Met 117	-	-	-	-	-	-
Glu 118	-	-	-	-	-	-
Leu 119	-	-	-	-	-	-
Tyr 120	0.10	0.25	0.26	0.29	-	-
Gly 121	0.05	0.04	0.04	0.04	0.06	0.09
Arg 122	-	0.03	0.03	0.04	0.04	0.05
Thr 123	0.02	0.01	0.01	0.02	0.00	0.00
Lys 124	0.02	0.01	0.02	0.03	0.03	0.01
Asp 125	0.02	0.00	0.01	0.02	0.02	0.01
Leu 126	0.02	0.01	0.02	0.03	0.03	-
Ser 127	0.01	0.01	0.01	0.02	0.02	0.01

Ser 128		0.01	0.01	0.02	0.02	0.00
Asp 129	0.01	0.01	0.03	0.01	0.02	0.00
Ile 130	0.03	0.02	0.02	0.01	0.01	0.03
Lys 131	0.02	0.02	0.02	0.02	0.02	0.02
Glu 132	0.03	0.02	0.03	0.05	0.04	0.03
Lys 133	0.03	0.00	0.01	0.02	0.03	0.02
Phe 134	0.04	0.03	0.03	0.03	0.00	0.02
Ala 135	0.03	0.05	0.04	0.02	0.02	0.06
Lys 136	0.02	0.02	0.02	0.04	0.04	0.01
Leu 137	0.05	0.02	0.01	-	0.02	-
Cys 138	0.02	0.02	0.02	-	0.01	0.05
Val 139	0.03	0.03	0.04	0.06	0.06	0.03
Ala 140	0.02	0.01	0.02	0.04	-	-
His 141	0.03	0.02	0.01	0.02	0.01	0.02
Gly 142	0.02	0.02	0.03	0.05	0.05	0.03
Ile 143	0.03	0.01	0.02	0.01	0.02	0.05
Thr 144	0.02	0.02	0.02	0.02	0.00	-
Arg 145	0.01	0.05	0.05	0.04	0.03	0.02
Asp 146	0.07	0.01	0.02	-	0.10	-
Asn 147	0.02	0.00	0.00	0.03	0.02	0.04
Ile 148	0.02	0.01	0.02	0.04	0.04	0.01
Ile 148	0.02	0.04	0.03	0.02	0.01	0.02
Asp 150	0.02	0.03	0.03	0.04	0.04	0.03
Leu 151	0.01	0.04	0.05	0.03	0.03	0.03
Thr 152	0.02	0.01	0.01	0.02	-	0.01
Lys 153	0.03	0.02	0.01	0.01	0.03	0.04
Thr 154	0.04	0.02	0.02	0.02	-	-
Asp 155	0.09	0.09	0.09	0.06	0.06	-
Arg 156	-	-	-	-	-	-
Cys 157	-	-	-	-	-	-
Leu 158	-	0.01	0.01	0.02	0.02	0.03
Gln 159	0.03	0.01	0.01	0.02	0.02	0.00
Ala 160	0.03	0.02	0.01	0.02	0.01	0.01

Appendix VIII $^1\text{H} / ^{15}\text{N}$ Heteronuclear NOE Values

Residue	UB NOE	NOE γ -hept	NOE γ -oct	NOE γ -non	NOE γ -dec	NOE γ -undec
Glu 1	0.39	0.38	0.38	0.33	0.30	0.41
Glu 2	0.44	0.39	0.36	0.32	0.42	0.43
Ala 3	0.43	0.40	0.34	0.40	0.45	0.44
Ser 4	0.48	0.54	0.38	0.66	0.58	0.40
Phe 5	0.66	0.72	0.70	0.72	0.67	0.64
Glu 6	0.58	-	-	-	-	0.58
Arg 7	0.63	0.51	0.50	0.66	0.60	0.47
Gly 8	0.54	0.50	0.53	0.64	0.53	0.52
Asn 9	0.61	0.59	0.62	0.47	0.58	0.61
Leu 10	0.57	0.00	0.59	0.85	0.80	0.78
Asp 11	0.81	0.81	0.87	0.52	0.87	0.68
Val 12	0.80	0.81	0.86	0.88	0.63	
Asp 13	0.73	0.81	0.71	0.72	0.85	0.91
Lys 14	0.81	0.72	0.80	1.04	0.76	0.82
Leu 15	0.80	0.87	0.80	0.71	0.79	0.81
Asn 16	0.84	0.86	0.82	0.85	0.79	0.87
Gly 17	0.90	0.73	0.89	0.68	0.52	0.81
Asp 18	0.82	0.73	0.84	0.67	0.82	0.49
Trp 19	0.83	0.92	0.80	0.80	0.93	0.90
Phe 20	0.91	0.92	0.76	0.34	0.96	0.94
Ser 21	0.71	0.70	0.79	1.07	1.19	0.86
Ile 22	0.83	-	0.75	0.73	0.85	0.81
Val 23	-	-	-	-	-	-
Val 24	0.80	0.93	0.79	1.12	0.88	-
Ala 25	0.45	0.80	0.86	0.83	0.92	0.74
Ser 26	0.80	0.78	0.89	0.98	0.81	0.83
Asp 27	0.91	1.02	0.85	0.95	0.63	-
Lys 28	-	-	-	-	-	-
Arg 29	0.50	0.81	0.75	0.93	0.75	0.76
Glu 30	0.77	0.73	0.82	0.71	0.78	0.74
Lys 31	0.88	0.85	0.85	0.76	0.84	0.79
Ile 32	0.85	1.01	0.88	1.41	0.82	0.84
Glu 33	0.77	0.83	0.79	0.79	0.80	0.78
Glu 34	0.77	0.83	0.82	0.85	0.75	-
Asn 35	0.79	0.70	0.84	0.52	0.99	0.79
Gly 36	0.81	0.75	0.93	0.99	0.68	0.74
Ser 37	0.69	0.81	0.99	0.85	0.77	0.74
Met 38	0.88	0.73	0.90	0.61	0.82	0.74
Arg 39	0.83	0.85	0.80	0.84	0.74	0.83
Val 40	0.81	0.83	0.82	0.71	0.71	0.91
Phe 41	-	-	-	-	-	-
Val 42	0.71	1.00	0.82	0.86	0.58	0.94

Gln 43	-	-	-	-	-	-
His 44	-	-	-	-	-	-
Ile 45	-	-	-	-	-	-
Asp 46	0.85	0.80	0.84	0.81	0.87	0.77
Val 47	0.86		1.12	0.76	0.84	0.79
Leu 48	0.74	0.83	0.80	0.82	0.70	0.74
Glu 49	0.59	0.58	0.54	0.59	0.55	0.63
Asn 50	0.90	0.71	0.64	0.41	0.71	0.69
Ser 51	0.75	0.80	0.74	0.85	0.77	0.73
Leu 52	0.84	0.80	0.85	1.02	0.97	0.88
Gly 53	0.93	0.81	0.75	0.97	1.20	0.79
Phe 54	0.72	0.72	0.85	-	0.96	-
Thr 55	0.85	0.80	0.86	-	-	-
Phe 56	0.87	0.73	0.82	-	0.82	0.82
Arg 57	0.88	0.82	0.89	0.80	0.87	0.95
Ile 58	0.82	0.82	0.76	0.80	0.70	0.81
Lys 59	0.82	0.79	0.80	0.75	1.06	0.84
Glu 60	0.58	0.63	0.73	0.77	0.55	0.62
Asn 61	0.68	0.60	0.68	0.75	0.65	0.63
Gly 62	0.65	0.69	0.77	0.70	0.59	0.74
Val 63	0.58	0.66	0.62	0.65	0.69	0.67
Cys 64	0.73	0.68	0.69	0.92	0.89	0.76
Thr 65	0.70	0.83	0.73	0.82	-	0.70
Glu 66	0.78	0.83	0.79	0.66	0.86	0.83
Phe 67	0.83	0.76	0.74	1.13	0.81	0.82
Ser 68	0.76	0.81	0.85	0.81	0.84	0.77
Leu 69	0.82	0.93	0.83	1.01	0.69	0.89
Val 70	0.78	0.77	0.75	0.86	0.67	1.00
Ala 71	0.83	0.80	0.82	0.68	0.72	-
Asp 72	0.71	0.78	0.69	0.64	0.63	0.77
Lys 73	0.54	-	-	-	-	-
Thr 74	0.62	0.76	0.70	0.63	-	-
Ala 75	0.77	-	-	-	-	-
Lys 76	0.74	0.74	0.81	0.70	0.96	0.76
Asp 77	0.72	0.67	0.69	0.63	0.74	0.37
Gly 78	0.77	0.61	0.61	0.95	0.22	0.75
Glu 79	0.71	0.87	0.75	0.92	0.74	0.81
Tyr 80	-	-	-	-	-	-
Phe 81	-	-	-	-	-	-
Val 82	0.82	0.93	0.77	0.72	0.75	0.79
Glu 83	0.78	0.73	0.78	0.67	0.87	0.73
Tyr 84	0.72	0.81	-	-	-	-
Asp 85	0.82	0.81	0.90	1.04	0.88	0.94

Gly 86	0.85	0.81	0.84	0.73	0.79	0.82
Glu 87	0.82	0.87	0.79	0.68	0.76	0.77
Asn 88	0.77	0.80	0.81	0.71	0.86	0.86
Thr 89	0.94	0.71	0.79	0.72	-	0.68
Phe 90	0.71	1.00	0.85	0.92	0.79	0.89
Thr 91	0.85	0.72	0.80	0.96	-	0.92
Ile 92	0.61	0.82	0.80	0.99	0.98	-
Leu 93	-	-	-	-	-	-
Lys 94	0.80	0.90	0.69	0.64	0.73	0.69
Thr 95	0.85	0.86	0.83	0.58	0.79	0.81
Asp 96	0.85	0.86	0.89	0.77	0.93	0.88
Tyr 97	-	1.18	0.75	-	0.93	1.01
Asp 98	0.73	0.83	0.89	0.98	0.85	0.77
Asn 99	0.91	0.89	0.75	0.96	0.87	0.88
Tyr 100	0.76	0.78	0.86	0.63	0.96	0.75
Val 101	0.87	-	-	-	-	-
Met 102	-	-	-	-	-	-
Phe 103	-	-	-	-	-	-
His 104	-	-	-	-	-	-
Leu 105	-	-	-	-	-	-
Val 106	1.32	0.81	0.75	1.24	0.79	0.78
Asn 107	-	0.95	0.77	0.59	0.82	0.78
Val 108	0.79	0.81	1.03	0.76	0.98	0.66
Asn 109	0.78	0.85	0.85	0.87	0.74	0.80
Asn 110	0.65	0.68	0.75	0.68	0.75	0.71
Gly 111	0.79	0.72	0.80	0.90	0.78	0.68
Glu 112	0.74	0.78	0.77	0.72	0.83	0.73
Thr 113	0.83	0.77	0.80	0.81	-	0.73
Phe 114	0.64	0.72	0.84	0.92	0.70	0.73
Gln 115	0.85	0.74	0.84	0.88	0.72	0.73
Leu 116	1.18	0.93	0.81	1.30	-	-
Met 117	-	-	-	-	-	-
Glu 118		-	-	-	-	-
Leu 119	-	-	-	-	-	-
Tyr 120	-	0.79	0.78	1.55	-	-
Gly 121	0.84	0.98	0.84	0.75	0.91	0.76
Arg 122	0.79	0.72	0.83	0.86	1.01	0.72
Thr 123	0.79	0.87	0.79	1.03		0.87
Lys 124	0.93	0.82	0.90	0.75	0.91	0.87
Asp 125	0.79	0.86	0.89	0.73	0.78	0.81
Leu 126	0.90	0.90	0.85	0.79	0.79	-
Ser 127	0.83	0.78	0.75	0.71	0.79	0.75
Ser 128	0.87	0.83	0.97	1.00	1.00	0.91
Asp 129	0.83	0.87	0.85	0.87	0.72	-
Ile 130	0.84	0.80	0.75	0.85	0.83	0.86

Lys 131	0.84	0.92	0.84	0.83	0.97	0.87
Glu 132	0.78	0.74	0.81	1.05	0.84	0.81
Lys 133	0.72	0.80	0.83	0.73	0.84	0.82
Phe 134	0.86	0.84	0.71	0.82	0.87	0.81
Ala 135	0.86	0.82	0.79	0.90	0.74	0.73
Lys 136	0.83	0.75	0.85	0.86	0.86	0.76
Leu 137	-	-	0.88	-	-	-
Cys 138	0.83	0.91	0.86	0.75	0.94	-
Val 139	0.90	1.67	0.73	0.97	0.64	0.71
Ala 140	0.72	0.88	0.78	0.86	-	-
His 141	0.86	0.73	0.84	0.72	0.88	0.69
Gly 142	0.81	0.82	0.81	0.84	0.60	0.76
Ile 143	0.80	0.84	0.78	0.82	0.80	0.82
Thr 144	0.78	0.67	0.68	0.87	-	-
Arg 145	0.78	0.77	0.84	0.90	0.85	0.89
Asp 146	0.80	0.79	0.86	-	-	-
Asn 147	0.87	0.82	0.85	0.96	0.82	0.76
Ile 148	0.85	0.82	0.85	0.85	0.83	0.83
Ile 148	0.89	0.90	0.90	0.80	0.71	0.84
Asp 150	0.75	0.82	0.83	0.80	0.83	0.85
Leu 151	0.66	0.80	0.86	0.73	1.23	0.88
Thr 152	0.90	1.14	0.77	0.78	-	0.75
Lys 153	0.67	0.51	0.67	0.55	0.67	0.72
Thr 154	0.79	0.80	0.74	0.73	-	-
Asp 155	0.83	0.96	0.81	0.72	0.84	-
Arg 156	-	-	-	-	-	-
Cys 157	-	-	-	-	-	-
Leu 158	0.80	0.77	0.76	0.92	0.80	0.65
Gln 159	0.37	0.36	0.42	0.52	0.41	0.36
Ala 160	-0.45	-0.46	-0.41	-0.48	-0.55	-0.42